(54) Title: RECEPTOR LIGAND VEGF-C

(57) Abstract

    Provided are polypeptide ligands for the receptor tyrosine kinase, Flt4. Also provided are cDNAs and vectors encoding the ligands, pharmaceutical compositions and diagnostic reagents comprising the ligands, and methods of making ans using the ligands.

- 1 -

RECEPTOR LIGAND VEGF-C

        This application is a continuation-in-part of
United States Patent Application Serial No. 08/671,573,
filed June 28, 1996, which is a continuation-in-part of
5   United States Patent Application Serial Number
08/601,132, filed February 14, 1996, which is a
continuation-in-part of United States Patent Application
Serial Number 08/585,895, filed January 12, 1996, which
is a continuation-in-part of United States Patent
10   Application Serial Number 08/510,133, filed August 1,
1995. This application also is a continuation-in-part of
United States Patent Application Serial Number
08/340,011, filed November 14, 1994.

FIELD OF THE INVENTION
15   The present invention generally relates to the
field of genetic engineering and more particularly to
growth factors for endothelial cells and growth factor
genes.

BACKGROUND OF THE INVENTION
20   Developmental growth, the remodelling and
regeneration of adult tissues, as well as solid tumor
growth, can only occur when accompanied by blood vessel
formation. Angioblasts and hematopoietic precursor cells
differentiate from the mesoderm and form the blood
25   islands of the yolk sac and the primary vascular system
of the embryo. The development of blood vessels from
these early (in situ) differentiating endothelial cells
is termed vasculogenesis. Major embryonic blood vessels
are believed to arise via vasculogenesis, whereas the
30   formation of the rest of the vascular tree is thought to
occur as a result of vascular sprouting from pre-existing
vessels, a process called angiogenesis, Risau et al.,
Devel. Biol., 125:441-450 (1988).
        Endothelial cells give rise to several types of
35   functionally and morphologically distinct vessels. When

- 2 -

organs differentiate and begin to perform their specific
functions, the phenotypic heterogeneity of endothelial
cells increases. Upon angiogenic stimulation,
endothelial cells may re-enter the cell cycle, migrate,
5   withdraw from the cell cycle and subsequently
differentiate again to form new vessels that are
functionally adapted to their tissue environment.
Endothelial cells undergoing angiogenesis degrade the
underlying basement membrane and migrate, forming
10 capillary sprouts that project into the perivascular
stroma. Ausprunk et al., Microvasc. Rev., 14:51-65
(1977). Angiogenesis during tissue development and
regeneration depends on the tightly controlled processes
of endothelial cell proliferation, migration,
15 differentiation, and survival. Dysfunction of the
endothelial cell regulatory system is a key feature of
many diseases. Most significantly, tumor growth and
metastasis have been shown to be angiogenesis dependent.
Folkman et al., J. Biol. Chem., 267:10931-10934 (1992).
20   Key signals regulating cell growth and
differentiation are mediated by polypeptide growth
factors and their transmembrane receptors, many of which
are tyrosine kinases. Autophosphorylated peptides within
the tyrosine kinase insert and carboxyl-terminal
25 sequences of activated receptors are commonly recognized
by kinase substrates involved in signal transduction for
the readjustment of gene expression in responding cells.
Several families of receptor tyrosine kinases have been
characterized. Van der Geer et al., Ann. Rev. Cell
30 Biol., 10:251-337 (1994). The major growth factors and
receptors transducing angiogenic stimuli are
schematically shown in Fig. 1.
        Fibroblast growth factors are also known to be
involved in the regulation of angiogenesis. They have
35 been shown to be mitogenic and chemotactic for cultured
endothelial cells. Fibroblast growth factors also
stimulate the production of proteases, such as

- 3 -

collagenases and plasminogen activators, and induce tube
formation by endothelial cells. Saksela *et al.*, *Ann.
Rev. Cell Biol.*, 4:93-126 (1988). There are two general
classes of fibroblast growth factors, FGF-1 and FGF-2,
5 both of which lack conventional signal peptides. Both
types have an affinity for heparin and FGF-2 is bound to
heparin sulfate proteoglycans in the subendothelial
extracellular matrix from which it may be released after
injury. Heparin potentiates the stimulation of
10 endothelial cell proliferation by angiogenic FGFs, both
by protecting against denaturation and degradation and
dimerizing the FGFs. Cultured endothelial cells express
the FGF-1 receptor but no significant levels of other
high-affinity fibroblast growth factor receptors.
15   Among other ligands for receptor tyrosine
kinases, the platelet derived growth factor, PDGF-BB, has
been shown to be weakly angiogenic in the chick
chorioallantoic membrane. Risau *et al.*, *Growth Factors*,
7:261-266 (1992). Transforming growth factor α (TGFα) is
20 an angiogenic factor secreted by several tumor cell types
and by macrophages. Hepatocyte growth factor (HGF), the
ligand of the c-met proto-oncogene-encoded receptor, also
is strongly angiogenic.
          Recent evidence shows that there are
25 endothelial cell specific growth factors and receptors
that may be primarily responsible for the stimulation of
endothelial cell growth, differentiation and certain
differentiated functions. The best studied of these is
vascular endothelial growth factor (VEGF), a member of
30 the PDGF family. Vascular endothelial growth factor is a
dimeric glycoprotein of disulfide-linked 23 kD subunits.
Other reported effects of VEGF include the mobilization
of intracellular calcium, the induction of plasminogen
activator and plasminogen activator inhibitor-1
35 synthesis, stimulation of hexose transport in endothelial
cells, and promotion of monocyte migration *in vitro*.
Four VEGF isoforms, encoded by distinct mRNA splice

variants, appear to be equally capable of stimulating
mitogenesis in endothelial cells. However, each isoform
has a different affinity for cell surface proteoglycans,
which behave as low affinity receptors for VEGF. The 121

5   and 165 amino acid isoforms of VEGF (VEGF121 and VEGF165)
are secreted in a soluble form, whereas the isoforms of
189 and 206 amino acid residues remain cell surface-
associated and have a strong affinity for heparin. VEGF
was originally purified from several sources on the basis

10 of its mitogenic activity toward endothelial cells, and
also by its ability to induce microvascular permeability,
hence it is also called vascular permeability factor
(VPF).

        The pattern of VEGF expression suggests its
15 involvement in the development and maintenance of the
normal vascular system and in tumor angiogenesis. During
murine development, the entire 7.5 day post-coital (p.c.)
endoderm expresses VEGF and the ventricular neuroectoderm
produces VEGF at the capillary ingrowth stage. See

20 Breier et al., Development, 114:521-523 (1992). On day
two of quail development, the vascularized area of the
yolk sac as well as the whole embryo show expression of
VEGF. In addition, epithelial cells next to fenestrated
endothelia in adult mice show persistent VEGF expression,

25 suggesting a role in the maintenance of this specific
endothelial phenotype and function.

        Two high affinity receptors for VEGF have been
characterized. These are VEGFR-1/Flt-1 (fms-like
tyrosine kinase-1) and VEGFR-2/Kdr/Flk-1 (kinase insert

30 domain containing receptor/fetal liver kinase-1). Those
receptors are classified in the PDGF-receptor family, but
they have seven rather than five immunoglobulin-like
loops in their extracellular domain (see Fig. 1) and they
possess a longer kinase insert than normally observed in

35 this family. The expression of VEGF receptors occurs
mainly in vascular endothelial cells, although some may
be present on hematopoietic progenitor cells, monocytes,

and melanoma cells. Only endothelial cells have been
reported to proliferate in response to VEGF, and
endothelial cells from different sources show different
responses. Thus, the signals mediated through VEGFR-1

5   and VEGFR-2 appear to be cell type specific. The VEGF-
related placenta growth factor (PlGF) was recently shown
to bind to VEGFR-1 with high affinity. PlGF was able to
enhance the growth factor activity of VEGF, but it did
not stimulate endothelial cells on its own. Naturally

10   occurring VEGF/PlGF heterodimers were nearly as potent
mitogens as VEGF homodimers for endothelial cells. Cao
et al., *J. Biol. Chem.*, *271*:3154-62 (1996).

The Flt4 receptor tyrosine kinase (VEGFR-3) is
closely related in structure to the products of the

15   VEGFR-1 and VEGFR-2 genes. Despite this similarity, the
mature form of Flt4 differs from the VEGF receptors in
that it is proteolytically cleaved in the extracellular
domain into two disulfide-linked polypeptides. Pajusola
et al., *Cancer Res.*, *52*:5738-5743 (1992). The 4.5 and

20   5.8 kb Flt4 mRNAs encode polypeptides which differ in
their C-termini due to the use of alternative 3' exons.
Isoforms of VEGF or PlGF do not show specific binding to
Flt4 or induce its autophosphorylation.

Expression of Flt4 appears to be more

25   restricted than the expression of VEGFR-1 or VEGFR-2.
The expression of Flt4 first becomes detectable by *in
situ* hybridization in the angioblasts of head mesenchyme,
the cardinal vein, and extraembryonically in the
allantois of 8.5 day p.c. mouse embryos. In 12.5 day

30   p.c. embryos, the Flt4 signal is observed in developing
venous and presumptive lymphatic endothelia, but arterial
endothelia appear negative. During later stages of
development, Flt4 mRNA becomes restricted to developing
lymphatic vessels. The lymphatic endothelia and some

35   high endothelial venules express Flt4 mRNA in adult human
tissues and increased expression occurs in lymphatic
sinuses in metastatic lymph nodes and in lymphangioma.

- 6 -

These results support the theory of the venous origin of
lymphatic vessels.

      Five endothelial cell specific receptor
tyrosine kinases, Flt-1 (VEGFR-1), KDR/Flk-1 (VEGFR-2),
5 Flt4 (VEGFR-3), Tie, and Tek/Tie-2 have so far been
described, which possess the intrinsic tyrosine kinase
activity essential for signal transduction. Targeted
mutations inactivating Flt-1, Flk-1, Tie, and Tek in
mouse embryos have indicated their essential and specific
10 roles in vasculogenesis and angiogenesis at the molecular
level. VEGFR-1 and VEGFR-2 bind VEGF with high affinity
(K_d 16 pM and 760 pM, respectively) and VEGFR-1 also binds
the related placenta growth factor (PlGF; K_d about 200
pM). A ligand for Tek is reported in PCT patent
15 publication WO 96/11269.

                    SUMMARY OF THE INVENTION
      The present invention provides a ligand,
designated VEGF-C, for the Flt4 receptor tyrosine kinase
(VEGFR-3). Thus, the invention provides a purified and
20 isolated polypeptide which is capable of binding to the
Flt4 receptor tyrosine kinase. Preferably, an Flt4
ligand of the invention is capable of stimulating
tyrosine phosphorylation of Flt4 receptor tyrosine kinase
in a host cell expressing the Flt4 receptor tyrosine
25 kinase. Preferred ligands of the invention are mammalian
polypeptides. Highly preferred ligands are human
polypeptides. As explained in detail below, dimers and
multimers comprising polypeptides of the invention linked
to each other or to other polypeptides are specifically
30 contemplated as ligands of the invention.
      In one embodiment, an FLT4 ligand polypeptide
has a molecular weight of approximately 23 kD as
determined by SDS-PAGE under reducing conditions. For
example, the invention includes a ligand composed of one
35 or more polypeptides of approximately 23 kD which is
purifyable from conditioned media from a PC-3 prostatic

- 7 -

adenocarcinoma cell line, the cell line having ATCC Acc.
No. CRL 1435. Amino acid sequencing of this PC-3 cell
derived ligand polypeptide revealed that the ligand
polypeptide comprises an amino terminal amino acid
5   sequence set forth in SEQ ID NO: 13. A conditioned
medium comprising an Flt4 ligand is itself an aspect of
the invention. The present invention also provides a new
use for the PC-3 prostatic adenocarcinoma cell line which
produces an Flt4 ligand. In a preferred embodiment, the
10 ligand may be purified and isolated directly from the PC-
3 cell culture medium.

In a highly preferred embodiment, the ligand
polypeptide comprises a fragment of the amino acid
sequence shown in SEQ ID NO: 33 which specifically binds
15 to the human Flt4 receptor tyrosine kinase. Exemplary
fragments include: a polypeptide comprising an amino acid
sequence set forth in SEQ ID NO: 33 from about residue
112 to about residue 213; a polypeptide comprising an
amino acid sequence from about residue 104 to about
20 residue 227 of SEQ ID NO: 33; and a polypeptide
comprising an amino acid sequence from about residue 112
to about residue 227 of SEQ ID NO: 33. Other exemplary
fragments include polypeptides comprising amino acid
sequences of SEQ ID NO: 33 that span, approximately, the
25 following residues: 31-213, 31-227, 32-227, 103-217,
103-225, 104-213, 113-213, 103-227, 113-227, 131-211,
161-211, 103-225, 227-419, 228-419, 31-419, and 1-419, as
described in greater detail below.

The present invention also provides one or more
30 polypeptide precursors of an Flt4 ligand, wherein one
such precursor (designated "prepro-VEGF-C") comprises the
complete amino acid sequence (amino acid residues 1 to
419) shown in SEQ ID NO: 33. Thus, the invention
includes a purified and isolated polypeptide having the
35 amino acid sequence of residues 1 to 419 shown in SEQ ID
NO: 33. Ligand precursors according to the invention,
when expressed in an appropriate host cell, produce, via

- 8 -

cleavage, a polypeptide which binds specifically to the
Flt4 receptor tyrosine kinase. A putative 102 amino acid
leader (prepro) peptide has been identified in the amino
acid sequence shown in SEQ ID NO: 33. Thus, in a
5   related aspect, the invention includes a purified and
isolated polypeptide having the amino acids sequence of
residues 103-419 shown in SEQ ID NO: 33.

In one embodiment, an expressed Flt4 ligand
polypeptide precursor is proteolytically cleaved upon
10   expression to produce an approximately 23 kD Flt4 ligand
polypeptide. Thus, an Flt4 ligand polypeptide is
provided which is the cleavage product of the precursor
polypeptide shown in SEQ ID NO: 33 and which has a
molecular weight of approximately 23 kD under reducing
15   conditions.

A putative VEGF-C precursor or splice variant,
consisting of polypeptides with molecular weights of
about 29 and 32 kD, also is considered an aspect of the
invention.

20   In another embodiment, an expressed Flt4 ligand
polypeptide precursor is proteolytically cleaved upon
expression to produce an approximately 21 kD VEGF-C
polypeptide. Sequence analysis has indicated that an
observed 21 kD form has an amino terminus approximately 9
25   amino acids downstream from the amino terminus of the 23
kD form, suggesting that alternative cleavage sites
exist.

From the foregoing, it will be apparent that an
aspect of the invention includes a fragment of the
30   purified and isolated polypeptide having the amino acid
sequence of residues 1 to 419 shown in SEQ ID NO: 33, the
fragment being capable of specifically binding to Flt4
receptor tyrosine kinase. Preferred embodiments include
fragments having an apparent molecular weight of
35   approximately 21/23 kD and 29/32 kD as assessed by SDS-
PAGE under reducing conditions.

Evidence suggests that the amino acids

- 9 -

essential for retaining Flt4 ligand activity are
contained within approximately amino acids 103/112-
226/227 of SEQ ID NO: 33, and that a carboxy-terminal
proteolytic cleavage to produce a mature, naturally-
5   occurring Flt4 ligand occurs at the approximate position
of amino acids 226-227 of SEQ ID NO: 33. Accordingly, a
preferred Flt4 ligand comprises approximately amino acids
103-227 of SEQ ID NO: 33.

VEGF-C mutational analysis described herein
10   indicates that a naturally occurring VEGF-C polypeptide
spanning amino acids 103-227 of SEQ ID NO: 33, produced
by a natural processing cleavage that defines the C-
terminus, exists and is biologically active as an Flt4
ligand. A polypeptide fragment consisting of residues
15   104-213 of SEQ ID NO: 33 has been shown to retain VEGF-C
biological activity. Additional mutational analyses
indicate that a polypeptide spanning only amino acids
113-213 of SEQ ID NO: 33 retains Flt4 ligand activity.
Accordingly, preferred polypeptides comprise sequences
20   spanning, approximately, amino acid residues 103-227,
104-213, or 113-213, of SEQ ID NO: 33.

Moreover, sequence comparisons of members of
the VEGF family of polypeptides provide an indication
that still smaller fragments will retain biological
25   activity, and such smaller fragments are intended as
aspects of the invention. In particular, eight highly
conserved cysteine residues of the VEGF family of
polypeptides define a region from residue 131 to residue
211 of SEQ ID NO: 33 (see Figures 10 & 31); therefore, a
30   polypeptide spanning from about residue 131 to about
residue 211 is expected to retain VEGF-C biological
activity. In fact, a polypeptide comprising
approximately residues 161-211, which retains an
evolutionarily-conserved RCXXCC motif, is postulated to
35   retain VEGF-C activity, and therefore is intended as an
aspect of the invention. Some of the conserved cysteine
residues in VEGF-C participate in interchain disulfide

- 10 -

bonding to make homo- and heterodimers of the various naturally occurring VEGF-C polypeptides. Beyond the preceding considerations, evidence exists that VEGF-C polypeptides lacking interchain disulfide bonds retain
5 VEGF-C biological activity. Consequently, the materials and methods of the invention include all VEGF-C fragments that retain at least one biological activity of VEGF-C, regardless of the presence or absence of interchain disulfide bonds. The invention also includes multimers
10 (including dimers) comprising such fragments linked to each other or to other polypeptides. Fragment linkage may be by way of covalent bonding (e.g., disulfide bonding) or non-covalent bonding of polypeptide chains (e.g, hydrogen bonding, bonding due to stable or induced
15 dipole-dipole interactions, bonding due to hydrophobic or hydrophilic interactions, combinations of these bonding mechanisms, and the like).

In yet another related aspect, the invention includes variants and analogs of the aforementioned
20 polypeptides, including VEGF-C, precursors of VEGF-C, and fragments of VEGF-C. The variants contemplated by the invention include purified and isolated polypeptides having amino acid sequences that differ from the amino acid sequences of VEGF-C, VEGF-C precursors and VEGF-C
25 fragments by conservative substitutions, as recognized by those of skill in the art, or by additions or deletions of amino acid residues that are compatible with the retention of at least one biological activity of VEGF-C.

Analogs contemplated by the invention include
30 polypeptides having modifications to one or more amino acid residues that differ from the modifications found in VEGF-C, VEGF-C precursors, or VEGF-C fragments, but are compatible with the retention of at least one biological activity of VEGF-C, VEGF-C precursors, or VEGF-C
35 fragments. For example, analogs within the scope of the invention include glycosylation variants and conjugants (attachment of the aforementioned polypeptides to

compounds such as labels, toxins, etc.)

The present invention also provides purified
and isolated polynucleotides (i.e., nucleic acids)
encoding polypeptides of the invention, for example a
5   cDNA or corresponding genomic DNA encoding a VEGF-C
precursor, VEGF-C, or biologically active fragments or
variants thereof. A preferred nucleic acid of the
invention comprises a DNA encoding amino acid residues 1
to 419 of SEQ ID NO: 33 or one of the aforementioned
10   fragments thereof. Due to the degeneracy of the genetic
code, numerous such coding sequences are possible, each
having in common the coding of the amino acid sequence
shown in SEQ ID NO: 33 or fragment thereof. The
invention also comprehends analogs of these
15   polynucleotides, or derivatives of any one of these
polynucleotides that encodes a polypeptide which retains
at least one biological activity of VEGF-C. DNA
polynucleotides according to the invention include
genomic DNAs, cDNAs, and oligonucleotides comprising the
20   coding sequence for a fragment of VEGF-C, or an analog of
a VEGF-C fragment that retains at least one of the
biological activities of VEGF-C. Distinct
polynucleotides encoding a polypeptide of the invention
by virtue of the degeneracy of the genetic code are
25   within the scope of the invention. In one embodiment,
the invention contemplates polynucleotides having
sequences that differ from polynucleotides encoding a
VEGF-C fragment in a manner that results in conservative
amino acid differences between the encoded polypeptides,
30   as understood by those of skill in the art.

A preferred polynucleotide according to the
invention comprises the human VEGF-C cDNA sequence set
forth in SEQ ID NO: 32 from nucleotide 352 to 1611.
Other polynucleotides according to the invention encode a
35   VEGF-C polypeptide from, e.g., mammals other than humans,
birds (e.g., avian quails), and others. Still other
polynucleotides of the invention comprise a coding

- 12 -

sequence for a VEGF-C fragment, and allelic variants of
those DNAs encoding part or all of VEGF-C.

The invention further comprises polynucleotides
that hybridize to the aforementioned polynucleotides
5   under standard stringent hybridization conditions.
Exemplary stringent hybridization conditions are as
follows: hybridization at 42°C in 50% formamide, 5X SSC,
20 mM Na·PO₄, pH 6.8 and washing in 0.2X SSC at 55°C. It
is understood by those of skill in the art that variation
10 in these conditions occurs based on the length and GC
nucleotide content of the sequences to be hybridized.
Formulas standard in the art are appropriate for
determining appropriate hybridization conditions. *See*
Sambrook *et al., Molecular Cloning: A Laboratory Manual*
15 (Second ed., Cold Spring Harbor Laboratory Press 1989) §§
9.47-9.51. These polynucleotides, capable of hybridizing
to polynucleotides encoding VEGF-C, VEGF-C fragments, or
VEGF-C analogs, are useful as nucleic acid probes for
identifying, purifying and isolating polynucleotides
20 encoding other (non-human) mammalian forms of VEGF-C.
Additionally, these polynucleotides are useful in
screening methods of the invention, as described below.

Preferred nucleic acid probes of the invention
comprise nucleic acid sequences of at least about 16
25 continuous nucleotides of SEQ ID NO: 32. More
preferably, these nucleic acid probes would have at least
about 20 nucleotides found in a subsequence of SEQ ID NO:
32. In using these nucleic acids as probes, it is
preferred that the nucleic acids specifically hybridize
30 to a portion of the sequence set forth in SEQ ID NO: 32.
Specific hybridization is herein defined as hybridization
under standard stringent hybridization conditions. To
identify and isolate other mammalian VEGF-C genes
specifically, nucleic acid probes preferably are selected
35 such that they fail to hybridize to genes related to
VEGF-C (*e.g.*, fail to hybridize to human VEGF or to human
VEGF-B genes).

Thus, the invention comprehends polynucleotides
comprising at least about 16 nucleotides wherein the
polynucleotides are capable of specifically hybridizing
to a gene encoding VEGF-C, e.g., a human gene. The
5   specificity of hybridization ensures that a
polynucleotide of the invention is able to hybridize to a
nucleic acid encoding a VEGF-C under hybridization
conditions that do not support hybridization of the
polynucleotide to nucleic acids encoding, e.g., VEGF or
10   VEGF-B. In one embodiment, polynucleotides of at least
about 16 nucleotides, and preferably at least about 20
nucleotides, are selected as continuous nucleotide
sequences found in SEQ ID NO: 32 or the complement of the
nucleotide sequence set forth in SEQ ID NO: 32.

15   Additional aspects of the invention include
vectors which comprise nucleic acids of the invention;
and host cells transformed or transfected with nucleic
acids or vectors of the invention. Preferred vectors of
the invention are expression vectors wherein nucleic
20   acids of the invention are operatively connected to
appropriate promoters and other control sequences, such
that appropriate host cells transformed or transfected
with the vectors are capable of expressing the Flt4
ligand. A preferred vector of the invention is plasmid
25   pFLT4-L, having ATCC accession no. 97231. Such vectors
and host cells are useful for recombinantly producing
VEGF-C polypeptides.

               The invention further includes a method of
making polypeptides of the invention. In a preferred
30   method, a nucleic acid or vector of the invention is
expressed in a host cell, and a polypeptide of the
invention is purified from the host cell or the host
cell's growth medium.

               In a related embodiment, the invention includes
35   a method of making a polypeptide capable of specifically
binding to Flt4 receptor tyrosine kinase, comprising the
steps of: (a) transforming or transfecting a host cell

- 14 -

with a nucleic acid of the invention; (b) cultivating the
host cell to express the nucleic acid; and (c) purifying
a polypeptide capable of specifically binding to Flt4
receptor tyrosine kinase from the host cell or from the
5   host cell's growth media.

The invention also is intended to include
purified and isolated polypeptide ligands of Flt4
produced by methods of the invention. In one preferred
embodiment, the invention includes a human VEGF-C
10   polypeptide or biologically active fragment thereof that
is substantially free of other human polypeptides.

In another aspect, the invention includes an
antibody which is specifically reactive with polypeptides
of the invention, or with polypeptides multimers of the
15   invention. Antibodies, both monoclonal and polyclonal,
may be made against a polypeptide of the invention
according to standard techniques in the art. Such
antibodies may be used in diagnostic applications to
monitor angiogenesis, vascularization, lymphatic vessels
20   and their disease states, wound healing, or certain tumor
cells, hematopoietic, or leukemia cells. The antibodies
also may be used to block the ligand from activating the
Flt4 receptor.

Ligands according to the invention may be
25   labeled with a detectable label and used to identify
their corresponding receptors *in situ*. Labeled Flt4
ligand and anti-Flt4 ligand antibodies may be used as
imaging agents in the detection of lymphatic vessels,
high endothelial venules and their disease states, and
30   Flt4 receptors expressed in histochemical tissue
sections. The ligand or antibody may be covalently or
non-covalently coupled to a suitable supermagnetic,
paramagnetic, electron dense, echogenic, or radioactive
agent for imaging. Other, non-radioactive labels, such
35   as biotin and avidin, may also be used.

A related aspect of the invention is a method
for the detection of specific cells, e.g., endothelial

- 15 -

cells. These cells may be found *in vivo*, or in *ex vivo*
biological tissue samples. The method of detection
comprises the steps of exposing a biological tissue
comprising, *e.g.*, endothelial cells, to a polypeptide
according to the invention which is capable of binding to
Flt4 receptor tyrosine kinase, under conditions wherein
the polypeptide binds to the cells, optionally washing
the biological tissue, and detecting the polypeptide
bound to the cells in the biological tissue, thereby
detecting the cells.

        The present invention also provides diagnostic
and clinical applications for claimed ligands. In a
preferred embodiment, Flt4 ligands or precursors are used
to accelerate angiogenesis, *e.g.*, during wound healing,
or to promote the endothelial functions of lymphatic
vessels. A utility for VEGF-C is suggested as an inducer
of angiogenesis also in tissue transplantation, in eye
diseases, in the formation of collateral vessels around
arterial stenoses and into injured tissues after
infarction. Polypeptides according to the invention may
be administered in any suitable manner using an
appropriate pharmaceutically-acceptable vehicle, *e.g.*, a
pharmaceutically-acceptable diluent, adjuvant, excipient
or carrier. VEGF-C polypeptides also may be used to
quantify future metastatic risk by assaying biopsy
material for the presence of active receptors or ligands
in a binding assay or kit using detectably-labeled VEGF-
C. An Flt4 ligand according to the invention also may be
used to promote re-growth or permeability of lymphatic
vessels in, for example, organ transplant patients. In
addition, an Flt4 ligand may be used to mitigate the loss
of axillary lymphatic vessels following surgical
interventions in the treatment of cancer (*e.g.*, breast
cancer). Ligands according to the invention also may be
used to treat or prevent inflammation, edema, aplasia of
the lymphatic vessels, lymphatic obstruction,
elephantiasis, and Milroy's disease. Finally, Flt4

- 16 -

ligands may be used to stimulate lymphocyte production
and maturation, and to promote or inhibit trafficking of
leukocytes between tissues and lymphatic vessels or to
affect migration in and out of the thymus.

5   . The invention includes a method of screening
for an endothelial cell disorder in a mammalian subject.
The method comprises providing a sample of endothelial
cells from the subject, contacting the sample of
endothelial cells with a polypeptide according to claim
10   4, determining the growth rate of the cells, and
correlating the growth rate with a disorder. In a
preferred embodiment, the endothelial cells are lymphatic
endothelial cells. In another preferred embodiment, the
mammalian subject is a human being and the endothelial
15   cells are human cells. In yet another preferred
embodiment, the disorder is a vessel disorder, e.g., a
lymphatic vessel disorder, such as the loss of lymphatic
vessels through surgery or the reduction in function of
existing lymphatic vessels due to blockages. In another
20   embodiment, the endothelial cells are contacted with the
polypeptide in vitro. The growth rate determined in the
method is the rate of cell division per unit time,
determined by any one of a number of techniques known in
the art. The correlation of the growth rate with a
25   disorder can involve a positive or negative correlation,
e.g., whether the polypeptide has Flt4 ligand activity or
is an inhibitor of such activity, as described below.

Inhibitors of the Flt4 ligand may be used to
control endothelial cell proliferation, lymphangiomas,
30   and metastatic cancer. For example, such inhibitors may
be used to arrest metastatic growth or spread, or to
control other aspects of endothelial cell expression and
growth. Inhibitors include antibodies, antisense
oligonucleotides, and polypeptides which block the Flt4
35   receptor, all of which are intended as aspects of the
invention.

In another embodiment, the invention provides a

method for modulating the growth of endothelial cells in
a mammalian subject comprising the steps of exposing
mammalian endothelial cells to a polypeptide according to
the invention in an amount effective to modulate the
growth of the mammalian endothelial cells. In one
embodiment, the modulation of growth is effected by using
a polypeptide capable of stimulating tyrosine
phosphorylation of Flt4 receptor tyrosine kinase in a
host cell expressing the Flt4 receptor tyrosine kinase.
In modulating the growth of endothelial cells, the
invention contemplates the modulation of endothelial
cell-related disorders. Endothelial cell disorders
contemplated by the invention include, but are not
limited to, physical loss of lymphatic vessels (e.g.,
surgical removal of axillary lymph tissue), lymphatic
vessel occlusion (e.g., elephantiasis), and
lymphangiomas. In a preferred embodiment, the subject,
and endothelial cells, are human. The endothelial cells
may be provided in vitro or in vivo, and they may be
contained in a tissue graft. An effective amount of a
polypeptide is defined herein as that amount of
polypeptide empirically determined to be necessary to
achieve a reproducible change in cell growth rate (as
determined by microscopic or macroscopic visualization
and estimation of cell doubling time, or nucleic acid
synthesis assays), as would be understood by one of
ordinary skill in the art.

         The present invention also provides methods for
using the claimed nucleic acids (i.e., polynucleotides)
in screening for endothelial cell disorders. In a
preferred embodiment, the invention provides a method for
screening an endothelial cell disorder in a mammalian
subject comprising the steps of providing a sample of
endothelial cell nucleic acids from the subject,
contacting the sample of endothelial cell nucleic acids
with a polynucleotide of the invention which is capable
of hybridizing to a gene encoding VEGF-C (and preferably

- 18 -

capable of hybridizing to VEGF-C mRNA), determining the
level of hybridization between the endothelial cell
nucleic acids and the polynucleotide, and correlating the
level of hybridization with a disorder. A preferred
mammalian subject, and source of endothelial cell nucleic
acids, is a human. The disorders contemplated by the
method of screening with polynucleotides include, but are
not limited to, vessel disorders such as the
aforementioned lymphatic vessel disorders, and hypoxia.

Purified and isolated polynucleotides encoding
other (non-human) VEGF-C forms also are aspects of the
invention, as are the polypeptides encoded thereby, and
antibodies that are specifically immunoreactive with the
non-human VEGF-C variants. Preferred non-human forms of
VEGF-C are forms derived from other vertebrate species,
including avian and mammalian species. Mammalian forms
are highly preferred. Thus, the invention includes a
purified and isolated mammalian VEGF-C polypeptide, and
also a purified and isolated polynucleotide encoding such
a polypeptide.

In one embodiment, the invention includes a
purified and isolated polypeptide having the amino acid
sequence of residues 1 to 415 of SEQ ID NO: 41, which
sequence corresponds to a putative mouse VEGF-C
precursor. The putative mouse VEGF-C precursor is
believed to be processed into a mature mouse VEGF-C in a
manner analogous to the processing of the human prepro-
polypeptide. Thus, in a related aspect, the invention
includes a purified and isolated polypeptide capable of
specifically binding to an Flt4 receptor tyrosine kinase
(e.g., a human or mouse Flt-4 receptor tyrosine kinase),
the polypeptide comprising a fragment of the purified and
isolated polypeptide having the amino acid sequence of
residues 1 to 415 of SEQ ID NO: 41, the fragment being
capable of specifically binding to the Flt4 receptor
tyrosine kinase. The invention further includes
multimers of the foregoing polypeptides and purified and

- 19 -

isolated nucleic acids encoding the foregoing
polypeptides, such as a nucleic acid comprising all or a
portion of the sequence shown in SEQ ID NO: 40.

        In another embodiment, the invention includes a
5   purified and isolated quail VEGF-C polypeptide,
biologically active fragments and multimers thereof, and
polynucleotides encoding the foregoing polypeptides.

        In yet another embodiment, the invention
includes a DNA comprising a VEGF-C promoter, that is
10 capable of promoting expression of a VEGF-C gene or
another operatively-linked, protein-encoding gene in
native host cells, under conditions wherein VEGF-C is
expressed in such cells.

                BRIEF DESCRIPTION OF THE DRAWING
15   Figure 1 schematically depicts major
endothelial cell receptor tyrosine kinases and growth
factors involved in vasculogenesis and angiogenesis.
Major structural domains are depicted, including
immunoglobulin-like domains (IGH), epidermal growth
20 factor homology domains (EGFH), fibronectin type III
domains (FNIII), transmembrane (TM) and juxtamembrane
(JM) domains, tyrosine kinase (TK1, TK2) domains, kinase
insert domains (KI), and carboxy-terminal domains (CT).

        Figure 2 schematically depicts the construction
25 of the pLTRFlt41 expression vector.

        Figure 3 schematically depicts the construction
of the baculovirus vector encoding a secreted soluble
Flt4 extracellular domain (Flt4EC).

        Figure 4 shows results of stimulation of Flt4
30 autophosphorylation by conditioned medium from PC-3 cell
cultures.

        Figures 5A, 5B, and 5C show that the major
tyrosyl phosphorylated polypeptide of Flt4-transfected
cells stimulated with PC-3 conditioned medium is the 125
35 kD Flt4 polypeptide (VEGFR-3), and also that the Flt4
stimulating activity is not adsorbed to heparin-

- 20 -

sepharose.

Figure 6 shows Western immunoblotting analysis of the Flt4 ligand activity isolated from PC-3 conditioned medium.

Figure 7 shows results of gel electrophoresis of chromatographic fractions from the affinity purification of Flt4 ligand (VEGF-C) isolated from concentrated PC-3 conditioned medium.

Figure 8 depicts the amino acid sequences of human, murine, and quail VEGF-C polypeptides, aligned to show similarity. Residues conserved in all three species are depicted in bold.

Figure 9 schematically depicts the cloning and structure of the Flt4 ligand, VEGF-C. The VEGF homologous region (dark shaded box) and amino and carboxyl terminal propeptides (light shaded and unshaded boxes, respectively) as well as putative signal sequence (ss) are depicted between 5' and 3' untranslated (ut) nucleic acid regions. The cleavage sites for the signal sequence and the amino and carboxyl terminal propeptides are indicated with triangles.

Figure 10 shows a comparison of the deduced amino acid sequences of PDGF-A, -B, PlGF-1, VEGF-B_{167}, VEGF165, and Flt4 ligand (VEGF-C).

Figure 11 depicts the exon-intron organization of the human VEGF-C gene. Seven exons are depicted as open boxes, with exon size depicted in base pairs. Introns are depicted as lines, with intron size (base pairs) depicted above the lines. 5' and 3' untranslated sequences of a putative 2.4 kb mature mRNA are depicted as shaded boxes. The location of genomic clones used to characterize the VEGF-C gene are depicted below the map of the gene.

Figure 12 shows Northern blotting analysis of the genes encoding VEGF, VEGF-B, AND VEGF-C (indicated by "FLT4-L") in two human tumor cell lines and in brain tissue.

- 21 -

Figure 13A is an autoradiograph showing recombinant VEGF-C isolated following a pulse-chase experiment and electrophoresed via SDS-PAGE under reducing conditions.

5   .Figure 13B is a photograph of polyacrylamide gel showing that recombinant VEGF-C forms are disulfide-linked in nonreducing conditions.

Figure 14A and 14B depict Western blots showing that VEGF-C stimulates autophosphorylation of VEGFR-2
10   (KDR) but has no effect on PDGFR-β phosphorylation.

Figure 15A shows that VEGF-C stimulates endothelial cell migration in a three-dimensional collagen gel assay.

Figure 16A shows the expression of VEGF-C mRNA
15   in human adult tissues.

Figure 16B shows the expression of VEGF, VEGF-B, and VEGF-C in selected human fetal tissues.

Figure 17 depicts the genomic structure of the human and murine VEGF-C genes. Sequences of exon-intron
20   junctions are depicted together with exon and intron lengths. Intron sequences are depicted in lower case letters. Nucleotides of the open reading frame observed in VEGF-C cDNAs are indicated as upper case letters in triplets (corresponding to the codons encoded at the
25   junctions).

Figure 18 presents a schematic illustration of VEGF-C processing, including the major forms of VEGF-C.

Figure 19 depicts autoradiograms from a pulse-chase immunoprecipitation experiment wherein cells
30   transfected with a VEGF-C expression vector (VEGF-C) and mock transfected cells (M) were pulse-labeled with radioactive amino acids and chased for varying lengths of time.

Figure 20 is a schematic map of the K14-VEGF-C
35   vector construct.

Figures 21A-C depict electrophoretic fractionations of the various forms of recombinant VEGF-C

- 22 -

produced by transfected 293 EBNA cells. Figure 21B
depicts the electrophoretic fractionation, under non-
reducing conditions, of polypeptides produced from mock
(M) transfected cells, cells transfected with wild type
5   (wt) VEGF-C cDNA, and cells transfected with a cDNA
variant encoding VEGF-C-R102S. Each of the bands
identified in Figure 21B was excised and
electrophoretically fractionated in a separate lane under
reducing conditions. Fractionation of bands
10   corresponding to wt VEGF-C are depicted in Figure 21A;
fractionation of bands corresponding to the R102S variant
are depicted in Figure 21C.

Figures 22A-B depict the forms and sizes of
wild type and mutant recombinant VEGF-Cs, as revealed by
15   non-reducing gel electrophoresis. Figure 22A shows the
VEGF-C forms secreted into the media; Figure 22B shows
the VEGF-C forms retained by the cells. Mock (M)
transfected cells served as a control.

Figures 23A-B present a comparison of the
20   pattern of immunoprecipitated, labelled VEGF-C forms
using antisera 882 and antisera 905. Adjacent lanes
contain immunoprecipitates that were (lanes marked +) or
were not (lanes marked -) subjected to reduction and
alkylation.

25   Figures 24A-B present Northern blots of total
RNA isolated from cells grown in the presence or absence
of interleukin-1 (IL-1) and/or dexamethasone (DEX) for
the indicated times. For Figure 24B, the Northern blot
was probed with radiolabeled DNA from a VEGF 581 bp cDNA
30   covering bps 57-638 (Genbank Acc. No. X15997), and a
human VEGF-B$_{167}$ cDNA fragment (nucleotides 1/382, Genbank
Acc. No. U48800). For Figure 24A, the Northern blot was
probed with radiolabeled DNA from a human full-length
VEGF-C cDNA (Genbank Acc. No. X94216). 18S and 28S- rRNA
35   markers.

Figure 25 shows VEGF-C expression in *P.
pastoris* cultures transfected with a VEGF-C cDNA, with

- 23 -

vector alone, or mock- (M) transfected, following
induction with methanol for various periods of time as
indicated. About 10 μl of medium was analyzed by gel
electrophoresis followed by Western blotting and
5   detection with anti-VEGF-C antiserum.

Figure 26 depicts the results of a Western blot
wherein NIH 3T3 cells expressing VEGFR-3 (Flt4), and PAE
cells expressing VEGFR-2 (KDR), were stimulated with 5x
concentrated medium from *Pichia* yeast transfected with a
10 VEGF-C cDNA-containing vector (+), with a vector lacking
an insert (-), or stimulated with the positive control
vanadate. The stimulated cells were lysed and
immunoprecipitated with VEGFR-specific antibodies, and
the immunoprecipitates were blotted and probed with anti-
15 phosphotyrosine antibodies.

Figures 27A-B present gel electrophoretograms
of human VEGF-C (wt) and VEGF-C variants secreted (Figure
27A) or retained (Figure 27B) by the host 293 EBNA cells.
Mock (M) transfected cells served as a control.
20 Molecular weight markers are indicated on the left in
kilodaltons (kD).

Figures 28A-B show Western blots of VEGFRs that
were stimulated to autophosphorylate by wild type (wt)
VEGF-C, as well as three VEGF-C polypeptide variants.
25 Cell lysates (NIH 3T3 for VEGFR-3 and PAE for VEGFR-2)
were subjected to receptor-specific antisera and the
receptors were immunoprecipitated. Immunoprecipitates
were then gel-fractionated and blotted for Western
analyses. Western blots were probed with anti-
30 phosphotyrosine antibodies.

Figures 29A-D are photomicrographs of
hematoxylin-eosin stained sections of K14-VEGF-C
transgenic and control mouse littermate tissues. Areas
shown are from the dorsal skin and snout, as indicated.
35 The white arrows show the endothelium-lined margin of the
lacunae devoid of red cells.

Figure 30 presents a Northern blot of

- 24 -

polyadenylated RNA from the indicated tissues, hybridized
with a pool of VEGF, VEGF-B$_{167}$, and VEGF-C probes.
Estimated transcript sizes are shown on the right in
kilobases (kb).

5   Figure 31 presents a comparison of the human
and mouse VEGF-C amino acid sequences. The amino acid
sequence of mouse VEGF-C is presented on the top line and
differences in the human sequence are marked below it.
The sequences have been labeled to depict the regions
10   shown in Figure 9. The arrow indicates the putative
cleavage site for the signal peptidase; BR3P motifs, as
well as a CR/SC motif, are boxed; and conserved cysteine
residues are marked in bold above the sequence. Arginine
residue 158 is also marked in bold. The numbering refers
15   to mouse VEGF-C residues.

           Figure 32 presents SDS-PAGE-fractionated
samples immunoprecipitated or affinity-purified from
various ^{35}S-labeled media. In the left panel, control
medium from Bosc23 cells containing vector only, medium
20   from cells expressing human VEGF-C, and medium from cells
expressing mouse VEGF-C were independently precipitated
with human VEGFR-3-Extracellular Domain coupled to
sepharose. In the right panel, similar conditioned media
were subjected to precipitation with anti-VEGF-C
25   antibodies. mwm: molecular weight markers; m- mouse; h-
human; α- anti.

           Figure 33 shows Western blots of gel-
fractionated immunoprecipitates from lysates made from
NIH 3T3 cells expressing VEGFR-3 or VEGFR-2, as
30   indicated, that had been stimulated by contact with VEGF-
C-containing lysates (or a vector control), as a measure
of VEGF-C-induced receptor autophosphorylation. Western
blots were probed with anti-phosphotyrosine (α-PTyr) or
anti-receptor antisera (anti-VEGFR-3 and anti-VEGFR-2),
35   as indicated. As a control, receptor autophosphorylation
was induced by pervanadate treatment (VO4). The arrows
and numbers refer to the apparent molecular weights of

- 25 -

the tyrosyl phosphorylated receptor polypeptide bands.
bVEGF: human baculoviral VEGF-C protein; C-FGEVm: lysate
from cells harboring a mouse VEGF-C cDNA cloned into the
vector in an antisense orientation.

Figures 34A-D depict photomicrographs of *in
situ* hybridizations revealing the expression of VEGF-C
and VEGF-B mRNAs in a parasagittal section of a 12.5 day
mouse embryo. Figure 34A: VEGF-C probe; j- jugular
veins, mn- metanephros, m- mesenterium (arrowheads), vc-
intervertebral vessels, lu- lung (arrowheads). Figure
34B: VEGF-B probe; h- heart, nasopharyngeal area
(arrowheads). Figure 34C: VEGF-C sense strand probe
serving as a control. Figure 34D: bright-field
photomicrograph of the same field shown in Figure 34C.

Figures 35A-H depict sections of mouse embryos
providing comparisons of VEGF-C and VEGFR-3 expression in
the jugular vessels and the mesenteric area. Figures 35A
and 35C show expression of VEGF-C transcripts in the
mesenchyme around the large sac-like structures in the
jugular area (arrowheads). Figures 35B and 35D show
expression of VEGFR-3 transcripts in the jugular venous
sacs. Figures 35E and 35G show VEGF-C mRNA distribution
in the mesenteric region of a 14.5 day p.c. embryo, as
well as around the gut. Figures 35F and 35H show VEGFR-3
mRNA in the mesenteric region of a 14.5 day embryo, as
well as the gut area, developing lymphatic vessels, and
venules.

Figures 36A-D depict photomicrographs showing FLT4
and VEGF-C *in situ* hybridization of the cephalic region
of a 16-day p.c. mouse embryo. A section of the cephalic
region hybridized with the Flt4 probe (Figure 36A) shows
the developing snout, nasal structures and eyes. A more
caudally located section shows hybridization with the
VEGF-C probe (Figure 36B). The round structures on both
sides in the upper part represent the developing molars.
In the upper (dorsal) part on both sides of the midline,
the caudal portion of the developing conchae are seen.

- 26 -

These structures also are shown in higher magnification darkfield (Figure 36C) and lightfield (Figure 36D) microscopy.

DETAILED DESCRIPTION OF THE INVENTION

5   Described herein is the isolation of a novel vascular endothelial growth factor and the cloning of a DNA encoding this novel growth factor from a cDNA library prepared from the human prostatic adenocarcinoma cell line PC-3. The isolated cDNA encodes a protein which is

10   proteolytically processed and secreted to cell culture medium. The secreted protein, designated VEGF-C, binds to the extracellular domain of Flt4 and induces tyrosine autophosphorylation of Flt4 and VEGFR-2. VEGF-C also stimulates the migration of endothelial cells in collagen

15   gel.

           Data reported herein indicates that VEGF-C is expressed as a larger precursor which is cleaved to produce the ligand. A coexpressed region in some cases results from alternative splicing of RNA of the ligand

20   gene. Such a co-expressed region may be a function of the particular expression system used to obtain the ligand. The skilled artisan understands that in recombinant production of proteins, additional sequence may be expressed along with a functional polypeptide

25   depending upon the particular recombinant construct used to express the protein, and subsequently removed to obtain the desired ligand. In some cases the recombinant ligand can be made lacking certain residues of the endogenous/natural ligand. Moreover, it is well-known in

30   that conservative replacements may be made in a protein which do not alter the function of the protein. Accordingly, it is anticipated that such alterations are within the scope of the invention. Moreover, it is anticipated that one or more VEGF-C precursors (the

35   largest putative native secreted VEGF-C precursor having the complete amino acid sequence from residue 32 to

- 27 -

residue 419 of SEQ ID NO: 33) is capable of stimulating
the Flt4 ligand without any further processing, in a
manner similar to that in which VEGF stimulates its
receptor in its unprocessed form after the secretion and
5   concomitant release of the signal sequence.

Results reported herein show that Flt4 (VEGFR-
3) transmits signals for VEGF-C. This conclusion is
based on the specific binding of VEGF-C to recombinant
Flt4EC (Flt4 extracellular domain) protein and the
10   induction of VEGFR-3 autophosphorylation by medium from
VEGF-C transfected cells. In contrast, neither VEGF nor
PlGF showed specific binding to VEGFR-3 or induced its
autophosphorylation.

As set forth in greater detail below, the
15   putative prepro-VEGF-C has a deduced molecular mass of
46,883; a putative prepro-VEGF-C processing intermediate
has an observed molecular weight of about 32 kD; and
mature VEGF-C isolated form conditioned media has a
molecular weight of about 23 kD as assessed by SDS-PAGE
20   under reducing conditions. A major part of the
difference in the observed molecular mass of the purified
and recombinant VEGF-C and the deduced molecular mass of
the prepro-VEGF-C encoded by the VEGF-C open reading
frame (ORF) is attributable to proteolytic removal of
25   sequences at the amino-terminal and carboxyl-terminal
regions of the prepro-VEGF-C polypeptide. However,
proteolytic cleavage of the putative 102 amino acid
leader sequence is not believed to account for the entire
difference between the deduced molecular mass of 46,883
30   and the observed mass of about 23 kD, because the deduced
molecular weight of a polypeptide consisting of amino
acids 103-419 of SEQ ID NO: 33 is 35,881 kD. Evidence
indicates that a portion of the observed difference in
molecular weights is attributable to proteolytic removal
35   of amino acid residues in both the amino and carboxyl
terminal regions of the VEGF-C precursor. Extrapolation
from studies of the structure of PDGF (Heldin *et al.*,

- 28 -

Growth Factors, 8:245-52 (1993)) suggests that the region
critical for receptor binding and activation by VEGF-C is
contained within amino acids residues 104-213, which are
found in the secreted form of the VEGF-C protein (*i.e.*,
5   the form lacking the putative prepro leader sequence and
some carboxyterminal sequences). The 23 kD polypeptide
binding VEGFR-3 is likely to represent the VEGF-
homologous domain. After biosynthesis, the nascent VEGF-
C polypeptide may be glycosylated at three putative N-
10   linked glycosylation sites identified in the deduced
VEGF-C amino acid sequence. Polypeptides containing
modifications, such as N-linked glycosylations, are
intended as aspects of the invention.

The carboxyl terminal amino acid sequences,
15   which increase the length of the VEGF-C polypeptide in
comparison with other ligands of this family, show a
pattern of spacing of cysteine residues reminiscent of
the Balbiani ring 3 protein (BR3P) sequence (Dignam *et
al.*, *Gene*, 88:133-40 (1990); Paulsson *et al.*, *J. Mol.
20   Biol.*, 211:331-49 (1990)). This novel C-terminal silk
protein-like structural motif of VEGF-C may fold into an
independent domain, which, on the basis of the
considerations above, is at least partially cleaved off
after biosynthesis. Interestingly, at least one cysteine
25   motif of the BR3P type is also found in the carboxyl
terminus of VEGF. In our experiments both the putative
precursor and cleaved ligand were detected in the cell
culture media, suggesting cleavage by cellular proteases.
The determination of amino-terminal and carboxy-terminal
30   sequences of VEGF-C isolates allows the identification of
the proteolytic processing sites. The generation of
antibodies against different parts of the pro-VEGF-C
molecule allows the exact determination of the precursor-
product relationship and ratio, their cellular
35   distribution, and the kinetics of processing and
secretion.

VEGF-C has a conserved pattern of eight

- 29 -

cysteine residues, which may participate in the formation
of intra- and interchain disulfide bonds, creating an
antiparallel dimeric biologically active molecule,
similar to PDGF. Mutational analysis of the cysteine
5   residues involved in the interchain disulfide bridges has
shown that, in contrast to PDGF, VEGF dimers need to be
held together by these covalent interactions in order to
maintain biological activity. Disulfide linking of the
VEGF-C polypeptide chains was evident in the analysis of
10   VEGF-C in nonreducing conditions, although recombinant
protein also contained ligand-active VEGF-C forms which
lacked disulfide bonds between the polypeptides.

VEGFR-3, which distinguishes between VEGF and
VEGF-C, is closely related in structure to VEGFR-1 and
15   VEGFR-2. Finnerty et al., Oncogene, 8:2293-98 (1993);
Galland et al., Oncogene, 8:1233-40 (1993); Pajusola et
al., Cancer Res., 52:5738-43 (1992). Besides VEGFR-3,
VEGFR-2 tyrosine kinase also is activated in response to
VEGF-C. VEGFR-2 mediated signals cause striking changes
20   in the morphology, actin reorganization and membrane
ruffling of porcine aortic endothelial cells
overexpressing this receptor. In these cells, VEGFR-2
also mediated ligand-induced chemotaxis and mitogenicity.
Waltenberger et al., J. Biol. Chem., 269:26988-95 (1994).
25   Similarly, the receptor chimera CSF-1R/VEGFR-3 was
mitogenic when ectopically expressed in NIH 3T3
fibroblastic cells, but not in porcine aortic endothelial
cells (Pajusola et al., 1994). Consistent with such
results, the bovine capillary endothelial (BCE) cells,
30   which express VEGFR-2 mRNA but very little or no VEGFR-1
or VEGFR-3 mRNAs, showed enhanced migration when
stimulated with VEGF-C. Light microscopy of the BCE cell
cultures in collagen gel also suggested that VEGF-C
stimulated the proliferation of these cells. The data
35   thus indicate that the VEGF ligands and receptors show a
great specificity in their signalling, which may be cell-
type-dependent.

- 30 -

The expression pattern of the VEGFR-3 (Kaipainen et al., Proc. Natl. Acad. Sci. (USA), 92:3566-70 (1995)) suggests that VEGF-C may function in the formation of the venous and lymphatic vascular systems during embryogenesis. Constitutive expression of VEGF-C in adult tissues shown herein further suggests that this gene product also is involved in the maintenance of the differentiated functions of the lymphatic and certain venous endothelia where VEGFR-3 is expressed (Kaipainen et al., 1995). Lymphatic capillaries do not have well-formed basal laminae and an interesting possibility remains that the silk-like BR3P motif is involved in producing a supramolecular structure which could regulate the availability of VEGF-C in tissues. However, as shown here, VEGF-C also activates VEGFR-2, which is abundant in proliferating endothelial cells of vascular sprouts and branching vessels of embryonic tissues, but not so abundant in adult tissues. Millauer et al., Nature, 367:576-78 (1993). These data have suggested that VEGFR-2 is a major regulator of vasculogenesis and angiogenesis. VEGF-C may thus have a unique effect on lymphatic endothelium and a more redundant function, shared with VEGF, in angiogenesis and possibly in regulating the permeability of several types of endothelia. Because VEGF-C stimulates VEGFR-2 and promotes endothelial migration, VEGF-C may be useful as an inducer of angiogenesis of blood and lymphatic vessels in wound healing, in tissue transplantation, in eye diseases, and in the formation of collateral vessels around arterial stenoses and into injured tissues after infarction.

Taken together, these results show an increased complexity of signalling in the vascular endothelium. They reinforce the concept that when organs differentiate and begin to perform their specific functions, the phenotypic heterogeneity of endothelial cells increases in several types of functionally and morphologically

distinct vessels. However, upon stimulation by suitable
angiogenic stimuli, endothelial cells can re-enter the
cell cycle, migrate, withdraw from the cell cycle and
subsequently differentiate again to form new vessels that
5   are functionally adapted to their tissue environment.
This process of angiogenesis, concurrent with tissue
development and regeneration, depends on the tightly
controlled balance between positive and negative signals
for endothelial cell proliferation, migration,
10   differentiation and survival.

Previously-identified growth factors promoting
angiogenesis include the fibroblast growth factors,
hepatocyte growth factor/scatter factor, PDGF and TGF-α.
(See e.g., Folkman, *Nature Med.*, 1:27-31 (1995); Friesel
15   et al., *FASEB J.*, 9:919-25 (1995); Mustonen et al., *J.
Cell. Biol.*, 129:895-98 (1995). However, VEGF has been
the only growth factor relatively specific for
endothelial cells. The newly identified factors VEGF-B
[Olofsson et al., *Proc. Natl. Acad. Sci.*, 93:2578-81
20   (1996)] and VEGF-C thus increase our understanding of the
complexity of the specific and redundant positive signals
for endothelial cells involved in vasculogenesis,
angiogenesis, permeability, and perhaps also other
endothelial functions. Expression studies using Northern
25   blotting show abundant VEGF-C expression in heart and
skeletal muscle; other tissues, such as placenta, ovary,
small intestine, thyroid gland, kidney, prostate, spleen,
testis and large intestine also express this gene.
Whereas PlGF is predominantly expressed in the placenta,
30   the expression patterns of VEGF, VEGF-B and VEGF-C
overlap in many tissues, which suggests that members of
the VEGF family may form heterodimers and interact to
exert their physiological functions.

Targeted mutagenesis leading to inactivation of
35   the VEGF receptor loci in the mouse genome has shown that
VEGFR-1 is necessary for the proper organization of
endothelial cells forming the vascular endothelium, while

VEGFR-2 is necessary for the generation of both
endothelial and hematopoietic cells. This suggests that
the four genes of the VEGF family can be targets for
mutations leading to vascular malformations or
5   cardiovascular diseases.

The following Examples illustrate preferred
embodiments of the invention, wherein the isolation,
characterization, and function of Flt4 ligands and
ligand-encoding nucleic acids according to the invention
10 are shown.

EXAMPLE 1
Production of pLTRFlt4l expression vector

Construction of the LTR-Flt4l vector, encoding
the long form of Flt4 receptor tyrosine kinase, is
15 schematically shown in Fig. 2. The full-length Flt4s
(Flt4 short form) cDNA (Genbank Accession No. X68203, SEQ
ID NO: 36) was assembled by first subcloning the S2.5
fragment, reported in Pajusola et al., Cancer Res.,
52:5738-5743 (1992), incorporated by reference herein,
20 containing base pairs 56-2534 of the Flt4s into the EcoRI
site of the pSP73 vector (Promega, Madison, WI).

Since cDNA libraries used for screening of Flt4
cDNAs did not contain the extreme 5' protein-coding
sequences, inverse PCR was used for the amplification of
25 the 5' end of Flt4 corresponding to the first 12 amino
acid residues (MQRGAALCLRLW). Poly(A)$^+$ RNA was isolated
from human HEL erythroleukemia cells and double-stranded
cDNA, were synthesized using an Amersham cDNA Synthesis
System Plus kit (Amersham Corp., Buckinghamshire, U.K.)
30 and a gene-specific primer: 5'-TGTCCTCGCTGTCCTTGTCT-3'
(SEQ ID NO: 1), which was located 195 bp downstream of
the 5' end of clone S2.5. Double-stranded cDNA was
treated with T4 DNA polymerase to blunt the ends and cDNA
was purified by filtration with Centricon 100 filters
35 (Amicon Inc., Beverly, MA). Circularization of the
blunt-ended cDNA was performed by ligation in a total

- 33 -

volume of 150 microliters. The reaction mixture
contained a standard ligation buffer, 5% PEG-8000, 1 mM
DTT and 8 U of T4 DNA ligase (New England Biolabs,
Beverly, MA). Ligation was carried out at 16°C for 16

5   hours. Fifteen microliters of this reaction mix were
used in a standard PCR reaction (100 μl total volume)
containing 100 ng of Flt4-specific primers introducing
SacI and PstI restriction sites, and 1 unit of Taq DNA
polymerase (Perkin Elmer Cetus). Two rounds of PCR were

10 performed using 33 cycles per round (denaturation at 95°C
for 1 minute, annealing at 55°C for 2 minutes, and
elongation at 72°C for 4 minutes). The PCR mixture was
treated sequentially with the SacI and PstI restriction
enzymes, and after purification with MagicPCR Preps

15 (Promega), DNA fragments were subcloned into the
pGEM3Zf(+) vector for sequencing (Promega). The sequence
corresponded to the 5' end of the Flt4s cDNA clone
deposited in the Genbank Database as Accession No.
X68203.

20   The sequence encoding the first 12 amino acid
residues was added to the expression construct by
ligating an SphI-digested PCR fragment amplified using
reverse transcription-PCR of poly(A)$^+$ RNA isolated from
the HEL cells. The forward primer had the following

25 sequence: 5'-ACATGCATGC CACCATGCAG CGGGGCGCCG CGCTGTGCCT
GCGACTGTGG CTCTGCCTGG GACTCCTGGA-3' (SEQ ID NO: 2) (SphI
site underlined, translational start codon marked in
bold). The translation start codon is immediately
downstream from an optimized Kozak consensus sequence.

30 (Kozak, Nucl. Acids Res., 15: 8125-8148 (1987).) The
reverse primer, 5'-ACATGCATGC CCCGCCGGT CATCC-3' (SEQ ID
NO: 3) (SphI site underlined), to the 5' end of the S2.5
fragment, thus replacing the unique SphI fragment of the
S2.5 plasmid. The resulting vector was digested with

35 EcoRI and ClaI and ligated to a 138 bp PCR fragment
amplified from the 0.6 kb EcoRI fragment (base pairs 3789
to 4416 in the Genbank X68203 sequence) which encodes the

3' end of Flt4s shown in Figure 1 of Pajusola et al.,
Cancer Res., 52:5738-5743 (1992), using the
oligonucleotides 5'-CGGAATTCCC CATGACCCCA AC-3' (SEQ ID
NO: 4) (forward primer, EcoRI site underlined) and 5'-

5   CCATCGATGG ATCCTACCTG AAGCCGCTTT CTT-3' (SEQ ID NO: 5)
(reverse primer, ClaI site underlined). The coding
domain was completed by ligation of the 1.2 kb EcoRI
fragment (base pairs 2535-3789 of the sequence found in
Gen Bank Acc. No. X68203) into the above construct. The

10   complete cDNA was subcloned as a HindIII-ClaI(blunted)
fragment (this ClaI site was also included in the 3'
primer used to construct the 3' end of the coding
sequence) to the pLTRpoly expression vector reported in
Mäkelä et al., Gene, 118: 293-294 (1992) (Genbank

15   accession number X60280, SEQ ID NO: 37), incorporated by
reference herein, using its HindIII-AccI(blunted)
restriction sites.

       The long form of Flt4 (Flt4l) was produced by
replacing the 3'-end of the short form as follows: The 3'

20   region of the Flt4l cDNA was PCR-amplified using a gene-
specific oligonucleotide (SEQ ID NO: 7, see below) and a
pGEM 3Z vector-specific (SP6 promoter) oligonucleotide
5'-ATTTAGGTGACACTATA-3' (SEQ ID NO: 6) as reverse and
forward primers, respectively. The template for PCR was

25   an Flt4l cDNA clone containing a 495 bp EcoRI fragment
extending downstream of the EcoRI site at nucleotide 3789
of the Genbank X68203 sequence (the sequence downstream
of this EcoRI site is deposited as the Flt4 long form 3'
sequence having Genbank accession number S66407 (SEQ ID

30   NO: 38)). The gene-specific oligonucleotide contains a
BamHI restriction site located right after the end of the
coding region and has the following sequence: 5'-
CCATCGATGGATCCCGATGCTGCTTAGTAGCTGT-3' (SEQ ID NO: 7)
(BamHI site is underlined). The PCR product was digested

35   with EcoRI and BamHI and transferred in frame to the
LTRFlt4s vector fragment from which the coding sequences
downstream of the EcoRI site at base pair 2535 (see

- 35 -

sequence X68203) had been removed by *EcoRI-BamHI*
digestion. The resulting clone was designated
pLTRFlt41.· Again, the coding domain was completed by
ligation of the 1.2 kb *EcoRI* fragment (base pairs 2535-
5   3789 of sequence X68203) back into the resulting
construct.

EXAMPLE 2

Production and analysis of Flt41 transfected cells

        NIH 3T3 cells (60 % confluent) were co-
10   transfected with 5 micrograms of the pLTRFlt41 construct
and 0.25 micrograms of the pSV2neo vector containing the
neomycin phosphotransferase gene (Southern *et al.*, *J.*
Mol. Appl. Genet., *1*:327 (1982)), using the DOTAP
liposome-based transfection reagents (Boehringer-
15   Mannheim, Mannheim, Germany). One day after
transfection, the cells were transferred into selection
media containing 0.5 mg/ml geneticin (GIBCO, Grand
Island, N.Y.). Colonies of geneticin-resistant cells
were isolated and analyzed for expression of the Flt4
20   proteins. Cells were lysed in boiling lysis buffer
containing 3.3% SDS 125 mM Tris, pH 6.8. Protein
concentrations of the samples were measured by the BCA
method (Pierce, Rockford, IL). About 50 micrograms of
protein from each lysate were analyzed for the presence
25   of Flt4 by 6% SDS-polyacrylamide gel electrophoresis
(SDS-PAGE) and immunoblotting using antisera against the
carboxyl terminus of Flt4. Signals on Western blots were
revealed using the ECL method (Amersham).

        For production of anti-Flt4 antiserum, the Flt4
30   cDNA fragment encoding the 40 carboxy-terminal amino acid
residues of the short form: NH2-PMTPTTYKG SVDNQTDSGM
VLASEEFEQI ESRHRQESGFR-COOH (SEQ ID NO: 8) was cloned as
a 657 bp *EcoRI*-fragment into the pGEX-1λT bacterial
expression vector (Pharmacia-LKB, Inc., Uppsala, Sweden)
35   in frame with the glutathione-S-transferase coding
region. The resultant GST-Flt4S fusion protein was

- 36 -

produced in *E. coli* and purified by affinity
chromatography using a glutathione-Sepharose 4B column.
The purified protein was lyophilized, dissolved in
phosphate-buffered saline (PBS), mixed with Freund's
5   adjuvant and used for immunization of rabbits at bi-
weekly intervals using methods standard in the art
(Harlow *et al.*, *Antibodies: A Laboratory Manual* (Cold
Spring Harbor Laboratory Press, 1988)). Antisera were
used, after the fourth booster immunization, for
10   immunoprecipitation of Flt4 from transfected cells. Cell
clones expressing Flt4 were also used for ligand
stimulation analysis.

EXAMPLE 3
Construction of a Flt4 EC baculovirus vector
15   **and expression and purification of its product**

The construction of an Flt4 extracellular
domain (EC) baculovirus vector is schematically depicted
in Fig. 3. The Flt4-encoding cDNA was prepared in both a
long form and a short form, each being incorporated in a
20   vector under control of the Moloney murine leukemia virus
LTR promoter. The nucleotide sequence of the short form
of the Flt4 receptor is available from the Genbank
database as Accession No. X68203 and the specific 3'
segment of the long form cDNA is available under GenBank
25   Accession No. S66407.

The ends of a cDNA segment encoding the Flt4
extracellular domain (EC) were modified as follows: The
3' end of the Flt4 cDNA (Genbank Accession Number X68203)
extracellular domain sequence was amplified using primer
30   1116 (5'-CTGGAGTCGACTTGGCGGACT-3'; SEQ ID NO: 9, *Sal*I
site underlined) and primer 1315 (5'-
CGCGGATCCCTAGTGATGGTGATGGTGATGTCTACCTTCGATCATGCTGCCCTTAT
CCTC-3'; (SEQ ID NO: 10, *Bam*HI site underlined). The
sequence at the 5' end of primer 1315 is not
35   complementary to the Flt4 coding region. Inspection of
the sequence that is complementary to this region of

- 37 -

primer 1315 reveals in a 5' to 3' order, a stop codon,
six contiguous histidine codons (for subsequent
chromatographic purification of the encoded polypeptide
using a Ni-NTA column; Qiagen, Hilden, Germany), and an
5   added BamHI site. The amplified fragment was digested
with SalI and BamHI and used to replace a unique SalI-
BamHI fragment in the LTRFlt4 vector shown in Fig. 3.
The SalI-BamHI fragment that was replaced encodes the
Flt4 transmembrane and cytoplasmic domains. The result
10   was a modified LTRFlt4 vector.
        The 5' end without the Flt4 signal sequence
encoding region was amplified by PCR using the primer
1335 (5'-CCCAAGCTTGGATCCAAGTGGCTACTCCATGACC-3'; (SEQ ID
NO: 11) the primer contains added HindIII (AAGCTT) and
15   BamHI (GGATCC) restriction sites, which are underlined).
The second primer used to amplify the region encoding the
Flt4 signal sequence was primer 1332 5'-
GTTGCCTGTGATGTGCACCA-3'; SEQ ID NO: 12). The amplified
fragment was digested with HindIII and SphI (the HindIII
20   site (AAGCTT) is underlined in primer 1335 and the SphI
site is within the amplified region of the Flt4 cDNA).
The resultant HindIII-SphI fragment was used to replace a
HindIII-SphI fragment in the modified LTRFlt4 vector
described immediately above (the HindIII site is in the
25   5' junction of the Flt4 insert with the pLTRpoly portion
of the vector, the SphI site is in the Flt4 cDNA). The
resultant Flt4EC insert was then ligated as a BamHI
fragment into the BamHI site in the pVTBac plasmid
described in Tessier et al., Gene 98:177-183 (1991),
30   incorporated herein by reference. The relative
orientation of the insert was confirmed by partial
sequencing so that the open reading frame of the signal
sequence-encoding portion of the vector was adjacent to,
and in frame with, the Flt4 coding region sequence. The
35   Flt4EC construct was transfected together with
baculovirus genomic DNA into SF-9 cells by lipofection.
Recombinant virus was purified, amplified and used for

- 38 -

infection of High-Five cells (Invitrogen, San Diego, CA)
using methods standard in the art. The Flt4
extracellular domain (Flt4EC) was purified from the
culture medium of the infected High-Five cells using Ni-
5   NTA affinity chromatography according to manufacturer's
instructions (Qiagen) for binding and elution of the
6xHis tag encoded in the COOH-terminus of the recombinant
Flt4 extracellular domain.

EXAMPLE 4

10   Isolation of an Flt4 Ligand from Conditioned Media
         A human Flt4 ligand according to the invention
was isolated from media conditioned by a PC-3 prostatic
adenocarcinoma cell line (ATCC CRL 1435) in Ham's F-12
Nutrient mixture (GIBCO) containing 7% fetal calf serum
15 (FCS). The cells were grown according to the supplier's
instructions. In order to prepare the conditioned media,
confluent PC-3 cells were cultured for 7 days in Ham's F-
12 Nutrient mixture (GIBCO) in the absence of fetal
bovine serum (FBS). Medium was then cleared by
20 centrifugation at 10,000 g for 20 minutes. The medium
was then screened to determine its ability to induce
tyrosine phosphorylation of Flt4 by exposure to NIH 3T3
cells which had been transfected with Flt4-encoding cDNA
using the pLTRFlt41 vector. For receptor stimulation,
25 experiments, subconfluent NIH 3T3 cells were starved
overnight in serum-free DMEM medium (GIBCO) containing
0.2% bovine serum albumin (BSA). The cells were
stimulated with the conditioned media for 5 minutes,
washed twice with cold PBS containing 100 micromolar
30 vanadate, and lysed in RIPA buffer (10 mM Tris pH 7.5, 50
mM NaCl, 0.5% sodium deoxycholate, 0.5% Nonidet P40 (BDH,
Poole, England), 0.1% SDS, 0.1 U/ml Aprotinin (Boehringer
Mannheim), 1 mM vanadate) for receptor
immunoprecipitation analysis. The lysates were
35 centrifuged for 20 minutes at 15,000 x g. The
supernatants were incubated for 2 hours on ice with 3

microliters of the antiserum against the Flt4 C-terminus
described in Example 2. See also Pajusola et al.,
Oncogene, 8:2931-2937 (1993), incorporated by reference
herein.

5   After a two hour incubation in the presence of
anti-Flt4 antiserum, protein A-Sepharose (Pharmacia) was
added and incubation was continued for 45 minutes with
rotation. The immunoprecipitates were washed three times
with the immunoprecipitation buffer and twice with 10 mM
10   Tris, pH 7.5, before analysis by SDS-PAGE. Polypeptides
were transferred to nitrocellulose and analyzed by
Western blotting using Flt4- or phosphotyrosine-specific
antisera and the ECL method (Amersham Corp.). Anti-
phosphotyrosine monoclonal antibodies (anti-PTyr; PY20)
15   were purchased from Transduction Laboratories (Lexington,
Kentucky). In some cases, the filters were restained
with a second antibody after stripping. The stripping of
the filters was done for 30 minutes at 50°C in 100 mM 2-
mercaptoethanol, 2% SDS, 62.5 mM Tris-HCl, pH 6.7, with
20   occasional agitation.

     As shown in Fig. 4, the PC-3 conditioned medium
(PC-3CM), stimulated tyrosine phosphorylation of a 125 kD
polypeptide when Flt4- expressing NIH 3T3 cells were
treated with the indicated preparations of media, lysed,
25   and the lysates were immunoprecipitated with anti-Flt4
antiserum followed by SDS-PAGE, Western blotting, and
staining using anti-PTyr antibodies. The resulting band
was weakly phosphorylated upon stimulation with
unconcentrated PC-3 conditioned medium (lane 2). The 125
30   kD band comigrated with the tyrosine phosphorylated,
processed form of the mature Flt4 from pervanadate-
treated cells (compare lanes 2 and 7 of Fig. 4, see also
Fig. 5A). Comigration was confirmed upon restaining with
anti-Flt4 antibodies as is also shown in Fig. 5A (panel
35   on the right). In order to show that the 125 kD
polypeptide is not a non-specific component of the
conditioned medium reactive with anti-phosphotyrosine

antibodies, 15 microliters of conditioned medium were
separated by SDS-PAGE, blotted on nitrocellulose, and the
blot was stained with anti-PTyr antibodies. No signal
was obtained (Fig. 5B). Also, unconditioned medium
5   failed to stimulate Flt4 phosphorylation, as shown in
Fig. 4, lane 1.

Fig. 5C shows a comparison of the effects of
PC-3 CM stimulation (+) on untransfected (lanes 4 and 5),
FGFR-4-transfected (lanes 8 and 9) and Flt4-transfected
10 NIH 3T3 cells (lanes 1-3, 6 and 7). These results
indicate that neither untransfected NIH 3T3 cells nor NIH
3T3 cells transfected with FGFR-4 showed tyrosine
phosphorylation of a protein of about 125 kD upon
stimulation with the conditioned medium from PC-3 cells.
15 Analysis of stimulation by PC-3 CM pretreated with
Heparin-Sepharose CL-6B (Pharmacia) for 2 hours at room
temperature (lane 3) showed that the Flt4 ligand does not
bind to heparin.

As shown in Fig. 4, lane 3, stimulating
20 activity was considerably increased when the PC-3
conditioned medium was concentrated four-fold using a
Centricon-10 concentrator (Amicon). Fig. 4, lane 4,
shows that pretreatment of the concentrated PC-3
conditioned medium with 50 microliters of the Flt4
25 extracellular domain coupled to CNBr-activated sepharose
CL-4B (Pharmacia; about 1mg of Flt4EC domain/ml sepharose
resin) completely abolished Flt4 tyrosine
phosphorylation. Similar pretreatment of the conditioned
medium with unsubstituted sepharose CL-4B did not affect
30 stimulatory activity, as shown in Fig. 4, lane 5. Also,
the flow through obtained after concentration, which
contained proteins of less than 10,000 molecular weight,
did not stimulate Flt4 phosphorylation, as shown in Fig.
4, lane 6.

35 In another experiment, a comparison of Flt4
autophosphorylation in transformed NIH 3T3 cells
expressing LTRFlt41 was conducted, using unconditioned

medium, medium from PC-3 cells expressing the Flt4
ligand, or unconditioned medium containing either 50
ng/ml of VEGF165 or 50 ng/ml of PlGF-1. The cells were
lysed, immunoprecipitated using anti-Flt4 antiserum and
5   analyzed by Western blotting using anti-phosphotyrosine
antibodies.   Only the PC-3 conditioned medium expressing
the Flt4 ligand (lane Flt-4L) stimulated Flt4
autophosphorylation.

          The foregoing data show that PC-3 cells produce
10 a ligand which binds to the extracellular domain of Flt4
and activates this receptor.

                         EXAMPLE 5
                 Purification of the Flt4 Ligand
          The ligand expressed by human PC-3 cells as
15 characterized in Example 4 was purified and isolated
using a recombinantly-produced Flt4 extracellular domain
(Flt4EC) in affinity chromatography.

          Two harvests of serum-free conditioned medium,
comprising a total of 8 liters, were collected from 500
20 confluent 15 cm diameter culture dishes containing
confluent layers of PC-3 cells. The conditioned medium
was clarified by centrifugation at 10,000 x g and
concentrated 80-fold using an Ultrasette Tangential Flow
Device (Filtron, Northborough, MA) with a 10 kD cutoff
25 Omega Ultrafiltration membrane according to the
manufacturer's instructions. Recombinant Flt4
extracellular domain was expressed in a recombinant
baculovirus cell system and purified by affinity
chromatography on Ni-agarose (Ni-NTA affinity column
30 obtained from Qiagen). The purified extracellular domain
was coupled to CNBr-activated Sepharose CL-4B at a
concentration of 5 mg/ml and used as an affinity matrix
for ligand affinity chromatography.

          Concentrated conditioned medium was incubated
35 with 2 ml of the recombinant Flt4 extracellular domain-
Sepharose affinity matrix in a rolling tube at room

- 42 -

temperature for 3 hours. All subsequent purification
steps were at +4 °C. The affinity matrix was then
transferred to a column with an inner diameter of 15 mm
and washed successively with 100 ml of PBS and 50 ml of
5   10 mM Na-phosphate buffer (pH 6.8). Bound material was
eluted step-wise with 100 mM glycine-HCl, successive 6 ml
elutions having pHs of 4.0, 2.4, and 1.9. Several 2 ml
fractions of the eluate were collected in tubes
containing 0.5 ml 1 M Na-phosphate (pH 8.0). Fractions
10   were mixed immediately and dialyzed in 1 mM Tris-HCl (pH
7.5). Aliquots of 75 µl each were analyzed for their
ability to stimulate tyrosine phosphorylation of Flt4.
The ultrafiltrate, 100 µl aliquots of the concentrated
conditioned medium before and after ligand affinity
15   chromatography, as well as 15-fold concentrated fractions
of material released from the Flt4 extracellular domain-
Sepharose matrix during the washings were also analyzed
for their ability to stimulate Flt4 tyrosine
phosphorylation.
20   As shown in Fig. 6, lane 3, the concentrated
conditioned medium induced prominent tyrosine
phosphorylation of Flt4 in transfected NIH 3T3 cells
overexpressing Flt4. This activity was not observed in
conditioned medium taken after medium was exposed to the
25   Flt4 Sepharose affinity matrix described above (Fig. 6,
lane 4). The specifically-bound Flt4-stimulating
material was retained on the affinity matrix after
washing in PBS, 10 mM Na-phosphate buffer (pH 6.8), and
at pH 4.0 (Fig. 6, lanes 5-7, respectively), and it was
30   eluted in the first two 2 ml aliquots at pH 2.4 (lanes 8
and 9). A further decrease of the pH of the elution
buffer did not cause release of additional Flt4-
stimulating material (Fig. 6, lane 11). Fig. 6, lane 1
depicts a control wherein Flt4-expressing cells were
35   treated with unconditioned medium; lane 2 depicts the
results following treatment of Flt4-expressing cells with
the ultrafiltrate fraction of conditioned medium

containing polypeptides of less than 10 kD molecular
weight.

Small aliquots of the chromatographic fractions
were concentrated in a SpeedVac concentrator (Savant,
Farmingdale, N.Y.) and subjected to SDS-PAGE under
reducing conditions with subsequent silver staining of
the gel, a standard technique in the art. As shown in
Fig. 7, the major polypeptide, having a molecular weight
of approximately 23 kD (reducing conditions), was
detected in the fractions containing Flt4 stimulating
activity (corresponding to lanes 8 and 9 in Fig. 6).
That polypeptide was not found in the other
chromatographic fractions. On the other hand, besides
these bands and a very faint band having a 32 kD
mobility, all other components detected in the two active
fractions were also distributed in the starting material
and in small amounts in the other washing and eluting
steps after their concentration. Similar results were
obtained in three independent affinity purifications,
indicating that the 23 kD polypeptide specifically binds
to Flt4 and induces tyrosine phosphorylation of Flt4.

Fractions containing the 23 kD polypeptide were
combined, dried in a SpeedVac concentrator and subjected
to SDS-PAGE in a 12.5% gel. The proteins from the gel
were then electroblotted to Immobilon-P (PVDF) transfer
membrane (Millipore, Marlborough, MA) and visualized by
staining of the blot with Coomassie Blue R-250. The
region containing only the stained 23 kD band was cut
from the blot and subjected to N-terminal amino acid
sequence analysis in a Prosite Protein Sequencing System
(Applied Biosystems, Foster City, CA). The data were
analyzed using a 610A Data Analysis System (Applied
Biosystems). Analysis revealed a single N-terminal
sequence of NH$_2$-XEETIKFAAAHYNTEILK-COOH (SEQ ID NO: 13).

- 44 -

EXAMPLE 6

Construction of PC-3 cell cDNA library in a
eukaryotic expression vector

Human poly(A)⁺ RNA was isolated from five 15 cm
5   diameter dishes of confluent PC-3 cells by a single step
method using oligo(dT) (Type III, Collaborative
Biomedical Products, Becton-Dickinson Labware, Bedford,
MA) cellulose affinity chromatography (Sambrook et al.,
1989). The yield was 70 micrograms. Six micrograms of
10   the Poly(A)⁺ RNA were used to prepare an oligo(dT)-primed
cDNA library in the mammalian expression vector pcDNA I
and the Librarian kit of Invitrogen according to the
instructions included in the kit. The library was
estimated to contain about 10⁶ independent recombinants
15   with an average insert size of approximately 1.8 kb.

EXAMPLE 7

Amplification of a unique nucleotide sequence
encoding the Flt4 ligand amino terminus

Degenerate oligonucleotides were designed based
20   on the N-terminal amino acid sequence of the isolated
human Flt4 ligand and were used as primers in a
polymerase chain reaction (PCR) to amplify cDNA encoding
the Flt4 ligand from the PC-3 cDNA library. The overall
strategy described in Examples 7 and 8 is schematically
25   depicted in Fig. 9, where the different primers have been
marked with arrows.

The PCR was carried out using 1 microgram of
DNA from the amplified PC-3 cDNA library and a mixture of
48 sense-strand primers present in equal proportions, the
30   primer sequences collectively comprising the sequence 5'-
GCAGARGARACNATHAA-3' (SEQ ID NO: 14) (wherein R is A or
G, N is A,G,C or T and H is A, C or T), encoding amino
acid residues 2-6 (EETIK, SEQ ID NO: 15) and 384
antisense-strand primers present in equal proportions,
35   the anti-sense strand primers collectively comprising the
sequence 5'-GCAYTTNARDATYTCNGT-3' (SEQ ID NO: 16)

(wherein Y is C or T and D is A, G or T), corresponding
to amino acid residues 14-18 (TEILK, SEQ ID NO: 17).
Three extra nucleotides (GCA) were added to the 5'-
terminus of each primer to increase annealing stability.
5   Two successive PCR runs were carried out using 1 U per
reaction of DynaZyme (F-500L, Finnzymes, Espoo, Finland),
a thermostable DNA polymerase, in a buffer supplied by
the manufacturer (10 mM Tris-HCl, pH 8.8 at 25°C, 1.5 mM
MgCl₂, 50 mM KCl, 0.1% Triton-X100), at an extension
10 temperature of 72°C. The first PCR run was carried out
for 43 cycles. The first three cycles were run at an
annealing temperature of 33°C for 2 minutes, and the
remaining cycles were run at 42°C for 1 minute.

        The region of the gel containing a weak band of
15 the expected size (57 bp) was cut out from the gel and
eluted. The eluted material was reamplified for 30
cycles using the same primer pairs described above at 42°C
for 1 minute. The amplified fragment was cloned into a
pCR II vector (Invitrogen) using the TA cloning kit
20 (Invitrogen) and sequenced using the radioactive
dideoxynucleotide sequencing method of Sanger. Six
clones were analyzed and all six clones contained the
sequence encoding the expected peptide (amino acid
residues 104-120 of the Flt4 ligand precursor).
25 Nucleotide sequence spanning the region from the third
nucleotide of codon 6 to the third nucleotide of codon 13
(the extension region) was identical in all six clones:
5'-ATTCGCTGCAGCACACTACAAC-3' (SEQ ID NO: 18) and thus
represented an amplified product from the unique sequence
30 encoding part of the amino terminus of the Flt4 ligand.

                        EXAMPLE 8

            Amplification of the 5'-end of the cDNA
                    encoding the Flt4 ligand

        Based on the unique nucleotide sequence
35 encoding the N-terminus of the isolated human Flt4
ligand, two pairs of nested primers were designed to

amplify, in two nested PCR reactions, the complete 5'-end
of the corresponding cDNAs from one microgram of DNA of
the above-described PC-3 cDNA library. First,
amplification was performed with an equal mixture of 4
5   primers collectively defining the sequence 5'-
TCNGTGTTGTAGTGTGCTG-3' (SEQ ID NO: 19), which is the
antisense-strand primer corresponding to amino acid
residues 9-15 (AAHYNTE, SEQ ID NO: 20), and sense-strand
primer 5'-TAATACGACTCACTATAGGG-3' (SEQ ID NO: 21),
10   corresponding to the T7 RNA promoter of the pcDNAI vector
used for construction of the library. "Touchdown" PCR
was used as disclosed in Don et al., Nucl. Acids Res.,
19:4008 (1991), incorporated by reference herein. The
annealing temperature of the two first cycles was 62°C and
15   subsequently the annealing temperature was decreased in
every other cycle by 1°C until a final temperature of 53°C
was reached, at which temperature 16 additional cycles
were conducted. Annealing time was 1 minute and
extension at each cycle was conducted at 72°C for 1
20   minute. Multiple amplified DNA fragments were obtained
in the first reaction. The products of the first
amplification (1 μl of a 1:100 dilution in water) were
used in the second amplification reaction employing a
pair of nested primers comprising an antisense-strand
25   primer 5'-GTTGTAGTGTGCTGCAGCGAATTT-3'; SEQ ID NO: 22)
encoding amino acid residues 6-13 (KFAAAHYN, SEQ ID NO:
23) of the Flt4 ligand, and a sense-strand primer (5'-
TCACTATAGGGAGACCCAAGC-3'; SEQ ID NO: 24), corresponding
to nucleotides 2179-2199 of the pcDNAI vector. The
30   sequences of these sense and antisense primers overlapped
with the 3' ends of the corresponding primers used in the
first PCR. "Touchdown" PCR was carried out by decreasing
the annealing temperature from 72°C to 66°C and continuing
with 18 additional cycles at 66°C. The annealing time was
35   1 minute and extension at each cycle was carried out at
72°C for 2 minutes. One major product of about 220 bp and
three minor products of about 270 bp, 150 bp, and 100 bp

were obtained.

The amplified fragment of approximately 220 bp
was excised from an agarose gel, cloned into a pCRII
vector using the TA cloning kit (Invitrogen), and

5   sequenced. Three recombinant clones were analyzed and
they contained the sequence 5'-
TCACTATAGGGAGACCCAAGCTTGGTACCGAGCTCGGATCCACTAGTAACGGCCGCC
AGTGTGGTGGAATTCGAC<u>GAACTCATGACTGTACTCTACCCAGAATATTGGAAAATG</u>
<u>TACAAGTGTCAGCTAAGGCAAGGAGGCTGGCAACATAACAGAGAACAGGCCAACCTC</u>

10   <u>AACTCAAGGACAGAAGAGACTATAAAATTCGCTGCAGCACACTACAAC</u>- 3' (SEQ
ID NO: 25). The beginning of the sequence represents the
pcDNAI vector and the underlined sequence represents the
amplified product of the 5'-end of the cDNA insert.

EXAMPLE 9

15   **Amplification of the 3'-end of
               cDNA encoding the Flt4 ligand**

Based upon the amplified 5'-sequence of the
clones encoding the amino terminus of the 23 kD human
Flt4 ligand, two pairs of non-overlapping nested primers

20   were designed to amplify the 3'-portion of the Flt-4-
ligand-encoding cDNA clones. The sense-strand primer 5'-
ACAGAGAACAGGCCAACC-3' (SEQ ID NO: 26), corresponding to
nucleotides 152-169 of the amplified 5'-sequences of the
Flt4 ligand (SEQ ID NO: 25), and antisense-strand primer

25   5'-TCTAGCATTTAGGTGACAC-3' (SEQ ID NO: 27) corresponding
to nucleotides 2311-2329 of the pcDNAI vector were used
in a first "touchdown" PCR. The annealing temperature of
the reaction was decreased 1˚C every two cycles from 72˚C
to 52˚C, at which temperature 15 additional cycles were

30   carried out. The annealing time was 1 minute and
extension at each cycle was carried out at 72˚C for 3
minutes. DNA fragments of several sizes were obtained in
the first amplification. Those products were diluted
1:200 in water and reamplified in PCR using the second

35   pair of primers: 5'-AAGAGACTATAAAATTCGCTGCAGC-3' (SEQ ID
NO: 28) and 5'-CCCTCTAGATGCATGCTCGA-3' (SEQ ID NO: 29)
(antisense-strand primer corresponding to nucleotides

- 48 -

2279-2298 of the pcDNAI vector). Two DNA fragments were obtained, having sizes of 1350 bp and 570 bp. Those fragments were cloned into a pCRII vector and the inserts of the clones were sequenced. Both of these fragments

5   were found to contain sequences encoding an amino acid sequence homologous to the VEGF sequence.

EXAMPLE 10

**Screening the PC-3 cell cDNA library using the 5'
PCR fragment of Flt4 ligand cDNA**

10   A 219 bp 5'-terminal fragment of human Flt4 ligand cDNA was amplified by PCR using the 5' PCR fragment described above and primers 5'-GTTGTAGTGTGCTGCAGCGAATTT-3' (antisense-strand primer, SEQ ID NO: 30) and 5'-TCACTATAGGGAGACCCAAGC-3' (SEQ ID NO:

15   31) (sense-primer corresponding to nucleotides 2179-2199 of the pcDNAI vector). The amplified product was subjected to digestion with *Eco*RI (Boehringer Mannheim) to remove the portion of the DNA sequence amplified from the pcDNAI vector and the resulting 153 bp fragment

20   encoding the 5' end of the Flt4 ligand was labeled with [^{32}P]-dCTP using the Klenow fragment of *E. coli* DNA polymerase I (Boehringer Mannheim). That fragment was used as a probe for hybridization screening of the amplified PC-3 cell cDNA library.

25   Filter replicas of the library were hybridized with the radioactively labeled probe at 42°C for 20 hours in a solution containing 50% formamide, 5x SSPE, 5x Denhardt's solution, 0.1% SDS and 0.1 mg/ml denatured salmon sperm DNA. Filters were washed twice in 1x SSC,

30   0.1% SDS for 30 minutes at room temperature, then twice for 30 minutes at 65°C and exposed overnight.

          On the basis of autoradiography, 10 positive recombinant bacterial colonies hybridizing with the probe were chosen from the library. Plasmid DNA was purified

35   from these colonies and analyzed by *Eco*RI and *Not*I digestion and agarose gel electrophoresis followed by

ethidium bromide staining. The ten plasmid clones were
divided into three groups on the basis of the presence of
insert sizes of approximately 1.7, 1.9 and 2.1 kb,
respectively. Inserts of plasmids from each group were

5   sequenced using the T7 oligonucleotide as a primer and
walking primers for subsequent sequencing reactions.
           Sequence analysis showed that all clones
contain the open reading frame encoding the NH2-terminal
sequence of the 23 kD human Flt4 ligand. Dideoxy

10   sequencing was continued using walking primers in the
downstream direction. A complete human cDNA sequence and
deduced amino acid sequence from a 2 kb clone is set
forth in SEQ ID NOs: 32 and 33, respectively. A putative
cleavage site of a "prepro" leader sequence is located

15   between residues 102 and 103 of SEQ ID NO: 33. When
compared with sequences in the GenBank Database, the
predicted protein product of this reading frame was found
to be homologous with the predicted amino acid sequences
of the PDGF/VEGF family of growth factors, as shown in

20   Fig. 10.
           Plasmid pFLT4-L, containing the 2.1 kb human
cDNA clone in pcDNAI vector, has been deposited with the
American Type Culture Collection, 12301 Parklawn Drive,
Rockville, MD 20852 as accession number 97231.

25   EXAMPLE 11

            Stimulation of Flt4 autophosphorylation by the
                protein product of the Flt4 ligand vector

           The 2.1 kb human cDNA insert of plasmid pFlt4-
L, which contains the open reading frame encoding the

30   sequence shown in SEQ ID NOs: 32 and 33; human VEGF-C,
see below), was cut out from the pcDNAI vector using
HindIII and NotI restriction enzymes, isolated from a
preparative agarose gel, and ligated to the corresponding
sites in the pREP7 expression vector (Invitrogen). The

35   pREP7 vector containing the pFlt4-L insert was
transfected into 293-EBNA cells (Invitrogen) using the

- 50 -

calcium phosphate transfection method (Sambrook et al.,
1989). About 48 hours after transfection the medium of
the transfected cells was changed to DMEM medium lacking
fetal calf serum and incubated for 36 h. The conditioned
5   medium was then collected, centrifuged at 5000 x g for 20
minutes, the supernatant was concentrated 5-fold using
Centriprep 10 (Amicon) and used to stimulate NIH 3T3
cells expressing LTRFlt41 (the Flt4 receptor), as in
Example 4. The cells were lysed, immunoprecipitated
10   using anti-Flt4 antiserum and analyzed by Western
blotting using anti-phosphotyrosine antibodies.

The conditioned medium from two different
dishes of the transfected cells stimulated Flt4
autophosphorylation in comparison with the medium from
15   mock-transfected cells, which gave only background levels
of phosphorylation of the Flt4 receptor. When the
concentrated conditioned medium was pre-absorbed with 20
microliters of a slurry of Flt4EC domain coupled to
Sepharose (see example 4), no phosphorylation was
20   obtained, showing that the activity responsible for Flt4
autophosphorylation was indeed the Flt4 ligand. Thus,
these results demonstrate that an expression vector
having an approximately 2.1 kb insert and containing an
open reading frame as shown in SEQ ID NO: 32 is expressed
25   as a biologically active Flt4 ligand (VEGF-C) in
transfected cells. The sequence encoded by that open
reading frame is shown in SEQ ID NO: 33.

The deduced molecular weight of a polypeptide
consisting of the complete amino acid sequence in SEQ ID
30   NO: 33 (residues 1 to 419) is 46,883. The deduced
molecular weight of a polypeptide consisting of amino
acid residues 103 to 419 of SEQ ID NO: 33 is 35,881. The
Flt4 ligand purified from PC-3 cultures had an observed
molecular weight of about 23 kD as assessed by SDS-PAGE
35   under reducing conditions. Thus, it appears that the
Flt4 ligand mRNA is translated into a precursor
polypeptide, from which the mature ligand is derived by

- 51 -

proteolytic cleavage. Also, the Flt4 ligand may be
glycosylated at three putative N-linked glycosylation
sites conforming to the consensus which can be identified
in the deduced Flt4 ligand amino acid sequence (N-
5   residues underlined in Fig. 10).

        The carboxyl terminal amino acid sequences,
which increase the predicted molecular weight of the Flt4
ligand subunit in comparison with other ligands of this
family, show a pattern of spacing of cysteine residues
10   reminiscent of the Balbiani ring 3 protein (BR3P)
sequence (Dignam et al., Gene, 88:133-140 (1990)), as
depicted schematically in Fig. 9. Such a sequence may
encode an independently folded domain present in a Flt4
ligand precursor and it may be involved, for example, in
15   the regulation of secretion, solubility, stability, cell
surface localization or activity of the Flt4 ligand.
Interestingly, at least one cysteine motif of the BR3P
type is also found in the VEGF carboxy terminal amino
acid sequences.

20   Thus, the Flt4 ligand mRNA appears first to be
translated into a precursor from the mRNA corresponding
to the cDNA insert of plasmid FLT4-L, from which the
mature ligand is derived by proteolytic cleavage. To
define the mature Flt4 ligand polypeptide, one first
25   expresses the cDNA clone (which is deposited in the
pcDNAI expression vector) in cells, such as COS cells.
One uses antibodies generated against encoded
polypeptides, fragments thereof, or bacterial Flt4 fusion
proteins, such as a GST-fusion protein, to raise
30   antibodies against the VEGF-homologous domain and the
amino- and carboxyl-terminal propeptides of Flt4 ligand.
One then follows the biosynthesis and processing of the
Flt4 ligand in the transfected cells by pulse-chase
analysis using radioactive cysteine for labelling of the
35   cells, immunoprecipitation and gel electrophoresis.
Using antibodies against the three domains of the product
encoded by the cDNA insert of plasmid FLT4-L, material

- 52 -

for radioactive or nonradioactive amino-terminal sequence
analysis is isolated. The determination of the amino-
terminal sequence of the mature VEGF-C polypeptide allows
for identification of the amino-terminal proteolytic

5   processing site. The determination of the amino-terminal
sequence of the carboxyl-terminal propeptide will give
the carboxyl-terminal processing site. This is
confirmed by site-directed mutagenesis of the amino acid
residues adjacent to the cleavage sites, which would

10   prevent the cleavage.

The Flt4 ligand is further characterizeable by
progressive 3' deletions in the 3' coding sequences of
the Flt4 ligand precursor clone, introducing a stop codon
resulting in carboxy-terminal truncations of its protein

15   product. The activities of such truncated forms are
assayed by, for example, studying Flt4
autophosphorylation induced by the truncated proteins
when applied to cultures of cells, such as NIH 3T3 cells
expressing LTRFlt4. By extrapolation from studies of the

20   structure of the related platelet derived growth factor
(PDGF, Heldin et al., Growth Factors, 8:245-252 (1993))
one determines that the region critical for receptor
activation by the Flt4 ligand is contained within the
first approximately 180 amino acid residues of the

25   secreted VEGF-C protein lacking the putative 102 amino
acid prepro leader (SEQ ID NO: 33, residues 103-282), and
apparently within the first approximately 120 amino acid
residues (SEQ ID NO: 33, residues 103-223).

On the other hand, the difference between the

30   molecular weights observed for the purified ligand and
deduced from the open reading frame of the Flt4 ligand
clone may be due to the fact that the soluble ligand was
produced from an alternatively spliced mRNA which would
also be present in the PC-3 cells, from which the

35   isolated ligand was derived. To isolate such alternative
cDNA clones one uses cDNA fragments of the deposited
clone and PCR primers made according to the sequence

- 53 -

provided as well as techniques standard in the art to
isolate or amplify alternative cDNAs from the PC-3 cell
cDNA library. One may also amplify using reverse
transcription (RT)-PCR directly from the PC-3 mRNA using
5   the primers provided in the sequence of the cDNA insert
of plasmid FLT4-L. Alternative cDNA sequences are
determined from the resulting cDNA clones. One can also
isolate genomic clones corresponding to the Flt4 ligand
mRNA transcript from a human genomic DNA library using
10 methods standard in the art and to sequence such clones
or their subcloned fragments to reveal the corresponding
exons. Alternative exons can then be identified by a
number of methods standard in the art, such as
heteroduplex analysis of cDNA and genomic DNA, which are
15 subsequently characterized.

EXAMPLE 12

Expression of the Gene Encoding
VEGF-C in Human Tumor Cell Lines

Expression of transcripts corresponding to the
20 Flt4 ligand (VEGF-C) was analyzed by hybridization of
Northern blots containing isolated poly(A)$^+$ RNA from HT-
1080 and PC-3 human tumor cell lines. The probe was the
radioactively labelled insert of the 2.1 kb cDNA clone
(pFlt4-L/VEGF-C, specific activity 10^8-10^9 cpm/mg of DNA).
25 The blot was hybridized overnight at 42°C using 50%
formamide, 5x SSPE buffer, 2% SDS, 10 x Denhardt's
solution, 100 mg/ml salmon sperm DNA and 1 x 10^6 cpm of
the labelled probe/ml. The blot was washed at room
temperature for 2 x 30 minutes in 2x SSC containing 0.05%
30 SDS, and then for 2 x 20 minutes at 52°C in 0.1x SSC
containing 0.1% SDS. The blot was then exposed at -70°C
for three days using intensifying screens and Kodak XAR
film. Both cell lines expressed an Flt4 ligand mRNA of
about 2.4 kb, as well as VEGF and VEGF-B mRNAs (Fig. 12).

‐ 54 ‐

EXAMPLE 13

VEGF-C Chains Are Proteolytically Processed
after Biosynthesis and Disulfide Linked

The predicted molecular mass of a secreted
5   human VEGF-C polypeptide, as deduced from the VEGF-C open
reading frame, is 46,883 kD, suggesting that VEGF-C mRNA
may be first translated into a precursor, from which the
ligands of 21/23 kD and 29/32 kD are derived by
proteolytic cleavage.

10   This possibility was explored by metabolic
labelling of 293 EBNA cells expressing VEGF-C.
Initially, 293 EBNA cells were transfected with the
VEGF-C construct. Expression products were labeled by
the addition of 100 μCi/ml of Pro-mix™ L-[^{35}S] in vitro
15   cell labelling mix ((containing ^{35}S-methionine and ^{35}S-
cysteine) Amersham, Buckinghamshire, England) to the
culture medium devoid of cysteine and methionine. After
two hours, the cell layers were washed twice with PBS and
the medium was then replaced with DMEM-0.2% BSA. After
20   1, 3, 6, 12 and 24 hours of subsequent incubation, the
culture medium was collected, clarified by
centrifugation, and concentrated, and human VEGF-C was
bound to 30 μl of a slurry of Flt4EC-Sepharose overnight
at +4°C, followed by three washes in PBS, two washes in 20
25   mM Tris-HCl (pH 7.5), alkylation, SDS-PAGE and
autoradiography. Alkylation was carried out by treatment
of the samples with 10mM 1,4 Dithiothreitol (Boehringer-
Mannheim, Mannheim, Germany) for one hour at 25°C, and
subsequently with 30 mM iodoacetamide (Fluka, Buchs,
30   Switzerland).

These experiments demonstrated that a putative
precursor polypeptide of 32 kD apparent molecular mass
was bound to the Flt4EC affinity matrix from the
conditioned medium of metabolically labelled cells
35   transfected with the human VEGF-C expression vector (Fig.
13A), but not from mock (M) transfected cells. Increased
amounts of a 23 kD receptor binding polypeptide

- 55 -

accumulated in the culture medium of VEGF-C transfected
cells during a subsequent chase period of three hours,
but not thereafter (lanes 2-4 and data not shown),
suggesting that the 23 kD form is produced by proteolytic
5   processing, which is incomplete, at least in the
transiently transfected cells. The arrows in Fig. 13A
indicate the 32 kD and 23 kD polypeptides of secreted
VEGF-C. Subsequent experiments showed that the 32kD
VEGF-C form contains two components migrating in the
10   absence of alkylation as polypeptides of 29 and 32 kD
(Figs. 21-23).

In a related experiment, human VEGF-C isolated
using Flt4EC-Sepharose after a 4 h continuous metabolic
labelling was analyzed by polyacrylamide gel
15   electrophoresis in nonreducing conditions (Fig. 13B).
Higher molecular mass forms were observed under
nonreducing conditions, suggesting that the VEGF-C
polypeptides can form disulfide-linked dimers and/or
multimers (arrows in Fig. 13B).

20   EXAMPLE 14
        Stimulation Of VEGFR-2 Autophosphorylation By VEGF-C
                Conditioned medium (CM) from 293 EBNA cells
transfected with the human VEGF-C vector also was used to
stimulate porcine aortic endothelial (PAE) cells
25   expressing VEGFR-2 (Kdr). Pajusola et al., Oncogene,
9:3545-55 (1994); Waltenberger et al., J. Biol. Chem.,
269:26988-26995 (1994). The cells were lysed and
immunoprecipitated using VEGFR-2 - specific antiserum
(Waltenberger et al., 1994).
30   PAE-KDR cells (Waltenberger et al., 1994) were
grown in Ham's F12 medium-10% fetal calf serum (FCS).
Confluent NIH 3T3-Flt4 cells or PAE-KDR cells were
starved overnight in DMEM or Ham's F12 medium,
respectively, supplemented with 0.2% bovine serum albumin
35   (BSA), and then incubated for 5 minutes with the analyzed
media. Recombinant human VEGF (R&D Systems) and PDGF-BB,

- 56 -

functional as stimulating agents, were used as controls.
The cells were washed twice with ice-cold Tris-Buffered
Saline (TBS) containing 100 mM sodium orthovanadate and
lysed in RIPA buffer containing 1 mM phenylmethylsulfonyl
5   fluoride (PMSF), 0.1 U/ml aprotinin and 1 mM sodium
orthovanadate. The lysates were sonicated, clarified by
centrifugation at 16,000 x g for 20 minutes and incubated
for 3-6 hours on ice with 3-5 μl of antisera specific for
Flt4 (Pajusola et al., 1993), VEGFR-2 or PDGFR-β
10   (Claesson-Welsh et al., J. Biol. Chem., 264:1742-1747
(1989); Waltenberger et al., 1994). Immunoprecipitates
were bound to protein A-Sepharose, washed three times
with RIPA buffer containing 1mM PMSF, 1mM sodium
orthovanadate, washed twice with 10 mM Tris-HCl (pH 7.4),
15   and subjected to SDS-PAGE using a 7% gel. Polypeptides
were transferred to nitrocellulose by Western blotting
and analyzed using PY20 phosphotyrosine-specific
monoclonal antibodies (Transduction Laboratories) or
receptor-specific antiserum and the ECL detection method
20   (Amersham Corp.).

The results of the experiment are presented in
Figs. 14A and 14B. As shown in Fig. 14A, PAE cells
expressing VEGFR-2 were stimulated with 10- or 2-fold
concentrated medium from mock-transfected 293-EBNA cells
25   (lanes 1 and 2), or with 2-, 5- or 10-fold concentrated
medium from 293-EBNA cell cultures expressing the
recombinant VEGF-C (lanes 3-6). VEGFR-2 was
immunoprecipitated with specific antibodies and analyzed
by SDS-PAGE and Western blotting using phosphotyrosine
30   antibodies. For comparison, the stimulations were
carried out with non-conditioned medium containing 50
ng/ml of purified recombinant VEGF (lanes 7 and 8).
Lanes 6 and 7 show stimulation with VEGF-C- or VEGF-
containing media pretreated with Flt4EC. As depicted in
35   Fig. 14B, PDGFR-β-expressing NIH 3T3 cells were
stimulated with non-conditioned medium (lane 1), 5-fold
concentrated CM from mock-transfected (lane 2) or VEGF-C

- 57 -

- transfected (lanes 3 and 4) cells, or with
non-conditioned medium containing 50 ng/ml of recombinant
human PDGF-BB (lane 5). Medium containing VEGF-C was
also pretreated with recombinant Flt4EC (lane 4). PDGFR-
5   β was immunoprecipitated with specific antibodies and
analyzed by SDS-PAGE and Western blotting using
phosphotyrosine antibodies with subsequent stripping and
reprobing of the membrane with antibodies specific for
PDGFR-β.

10   Referring again to Fig. 14A, a basal level of
tyrosine phosphorylation of VEGFR-2 was detected in cells
stimulated by CM from the mock-transfected cells. A
further concentration of this medium resulted in only a
slight enhancement of VEGFR-2 phosphorylation (lanes 1
15   and 2). CM containing recombinant VEGF-C stimulated
tyrosine autophosphorylation of VEGFR-2 and the intensity
of the autophosphorylated polypeptide band was increased
upon concentration of the VEGF-C CM (lanes 3-5).
Furthermore, the stimulating effect was abolished after
20   pretreatment of the medium with the Flt4EC affinity
matrix (compare lanes 1, 5 and 6). The maximal effect of
VEGF-C in this assay was comparable to the effect of
recombinant VEGF added to unconditioned medium at
concentration of 50 ng/ml (lane 8). Pretreatment of the
25   medium containing VEGF with Flt4EC did not abolish its
stimulating effect on VEGFR-2 (compare lanes 7 and 8).
These results suggest that the VEGF-C expression vector
encodes a ligand not only for Flt4 (VEGFR-3), but also
for VEGFR-2 (Kdr).

30   In order to further confirm that the
stimulating effect of VEGF-C on tyrosine phosphorylation
of VEGFR-3 and VEGFR-2 was receptor-specific, we analyzed
the effect of VEGF-C on tyrosine phosphorylation of PDGF
receptor β (PDGFR-β) which is abundantly expressed on
35   fibroblastic cells. As can be seen from Fig. 14B, a weak
tyrosine phosphorylation of PDGFR-β was detected upon
stimulation of Flt4-expressing NIH 3T3 cells with CM from

the mock-transfected cells (compare lanes 1 and 2). A
similar low level of PDGFR-β phosphorylation was observed
when the cells were incubated with CM from the VEGF-C
transfected cells, with or without prior treatment with
Flt4EC (lanes 3 and 4). In contrast, the addition of 50
ng/ml of PDGF-BB induced a prominent tyrosine
autophosphorylation of PDGFR-β (lane 5).

EXAMPLE 15

VEGF-C Stimulates Endothelial Cell
Migration In Collagen Gel

Conditioned media (CM) from cell cultures
transfected with the VEGF-C expression vector was placed
in a well made in collagen gel and used to stimulate the
migration of bovine capillary endothelial (BCE) cells in
the three-dimensional collagen gel as follows.

BCE cells (Folkman et al., Proc. Natl. Acad.
Sci. (USA), 76:5217-5221 (1979) were cultured as
described in Pertovaara et al., J. Biol. Chem., 269:6271-
74 (1994). The collagen gels were prepared by mixing
type I collagen stock solution (5 mg/ml in 1 mM HCl) with
an equal volume of 2x MEM and 2 volumes of MEM containing
10% newborn calf serum to give a final collagen
concentration of 1.25 mg/ml. The tissue culture plates
(5 cm diameter) were coated with about 1 mm thick layer
of the solution, which was allowed to polymerize at 37°C.
BCE cells were seeded on top of this layer. For the
migration assays, the cells were allowed to attach inside
a plastic ring (1 cm diameter) placed on top of the first
collagen layer. After 30 minutes, the ring was removed
and unattached cells were rinsed away. A second layer of
collagen and a layer of growth medium (5% newborn calf
serum (NCS)), solidified by 0.75% low melting point agar
(FMC BioProducts, Rockland, ME), were added. A well (3
mm diameter) was punched through all the layers on both
sides of the cell spot at a distance of 4 mm, and the
sample or control media were pipetted daily into the

- 59 -

wells. Photomicrographs of the cells migrating out from
the spot edge were taken after six days through an
Olympus CK 2 inverted microscope equipped with
phase-contrast optics. The migrating cells were counted
5   after nuclear staining with the fluorescent dye
bisbenzimide (1 mg/ml, Hoechst 33258, Sigma).

     Fig. 15 depicts a comparison of the number of
cells migrating at different distances from the original
area of attachment towards wells containing media
10   conditioned by the non-transfected (control) or
transfected (mock; VEGF-C; VEGF) cells, 6 days after
addition of the media. The number of cells migrating out
from the original ring of attachment was counted in five
adjacent 0.5 mm x 0.5 mm squares using a microscope
15   ocular lens grid and 10x magnification with a
fluorescence microscope. Cells migrating further than
0.5 mm were counted in a similar way by moving the grid
in 0.5 mm steps. The experiments were carried out twice
with similar results, and medium values from the one of
20   the experiments are presented with standard error bars.
As can be seen from the columns, VEGF-C-containing CM
stimulated cell migration more than medium conditioned by
the non-transfected or mock-transfected cells but less
than medium from cells transfected with a VEGF expression
25   vector. Daily addition of 1 ng of FGF2 into the wells
resulted in the migration of approximately twice the
number of cells when compared to the stimulation by CM
from VEGF-transfected cells.

                      EXAMPLE 16
30   **VEGF-C Is Expressed In Multiple Tissues**
          Northern blots containing 2 micrograms of
isolated poly(A)$^+$ RNA from multiple human tissues (blot
from Clontech Laboratories, Inc., Palo Alto, CA) were
probed with radioactively labelled insert of the 2.1 kb
35   VEGF-C cDNA clone. Northern blotting and hybridization
analysis showed that the 2.4 kb RNA and smaller amounts

- 60 -

of a 2.0 kb mRNA are expressed in multiple human tissues, most prominently in the heart, placenta, muscle, ovary and small intestine (Fig. 16A). Very little VEGF-C RNA was seen in the brain, liver or thymus and peripheral
5   blood leukocytes (PBL) appeared negative. A similar analysis of RNA from human fetal tissues (Fig. 16B) shows that VEGF-C is highly expressed in the kidney and lung and to a lesser degree in the liver, while essentially no expression is detected in the brain. Interestingly, VEGF
10   expression correlates with VEGF-C expression in these tissues, whereas VEGF-B is highly expressed in all tissues analyzed.

EXAMPLE 17
The VEGF-C Gene Localizes To Chromosome 4q34
15   A DNA panel of 24 interspecies somatic cell hybrids, which had retained one or two human chromosomes, was used for the chromosomal localization of the VEGF-C gene (Bios Laboratories, Inc., New Haven, CT). Primers were designed to amplify an about 250 bp fragment of the
20   VEGF-C gene from somatic cell hybrid DNA. The primers and conditions for polymerase chain reaction (PCR) were 5'-TGAGTGATTTGTAGCTGCTGTG-3' (forward) [SEQ ID NO: 34] and 5'-TATTGCAGCAACCCCCACATCT-3' (reverse) [SEQ ID NO: 35] for VEGF-C (94°C, 60s/62°C, 45s/72°C, 60s). The PCR
25   products were evaluated by electrophoresis in 1% agarose gels and visualized by ethidium bromide staining in ultraviolet light. [α-^{32}P]-dCTP- labelled cDNA inserts of a plasmid representing the complete VEGF-C coding domain was used as a probe in Southern blotting and
30   hybridization analysis of the somatic cell hybrid DNAs as instructed by the supplier (Bios Laboratories).
              The cell lines for fluorescence *in situ* hybridization (FISH) were obtained from the American Type Culture Collection (Rockville, MD). Purified DNA from P1
35   clones 7660 and 7661 (VEGF-C) (Genome Systems, Inc., St. Louis, MO) were confirmed positive by Southern blotting

of *EcoRI*- digested DNA followed by hybridization with the
VEGF-C cDNA. The P1 clones were then labelled by nick
translation either with biotin-11-dUTP, biotin-14-ATP
(Sigma Chemical Co., St. Louis, MO) or digoxigenin

5   11-dUTP (Boehringer Mannheim GmbH, Mannheim, Germany)
according to standard protocols. PHA-stimulated
peripheral blood lymphocyte cultures were treated with
5-bromodeoxyuridine (BrdU) at an early replicating phase
to induce G-banding. See Takahashi *et al.*, *Human Genet.*,

10 *86*:14-16 (1995); Lemieux *et al.*, *Cytogenet. Cell Genet.*,
59:311-12 (1992). The FISH procedure was carried out in
50% formamide, 10% dextran sulphate in 2x SSC using well-
known procedures. *See e.g.*, Rytkönnen *et al.*, *Cytogenet
Cell Genet.*, *68*:61-63 (1995); Lichter *et al.*, *Proc. Natl.*

15 *Acad. Sci. (USA)*, *85*:9664-68 (1988). Repetitive
sequences were suppressed with 50-fold excess of Cot-1
DNA (BRL, Gaithersburg, MD) compared with the labeled
probe. Specific hybridization signals were detected by
incubating the hybridized slides in labelled

20 antidigoxigenin antibodies, followed by counterstaining
with 0.1mmol/L 4,6-diamino-2-phenylindole. Probe
detection for two-color experiments was accomplished by
incubating the slides in fluorescein isothiocyanate
(FITC)-anti-digoxigenin antibodies (Sigma Chemical Co.)

25 and Texas red-avidin (Vector Laboratories, Burlingame,
CA) or rhodamine-anti-digoxigenin and FITC-avidin.
        Multi-color digital image analysis was used for
acquisition, display and quantification of hybridization
signals of metaphase chromosomes. The system contains a

30 PXL camera (Photometrics Inc., Tucson, AZ) attached to a
PowerMac 7100/Av workstation. IPLab software controls
the camera operation, image acquisition and Ludl Filter
wheel. At least 50 nuclei were scored. Overlapping
nuclei and clusters of cells were ignored. A slide

35 containing normal lymphocyte metaphase spreads and
interphase nuclei was included in each experiment to
control for the efficiency and specificity of the

- 62 -

hybridization.

In order to determine the chromosomal localization of the human VEGF-C gene, DNAs from human rodent somatic cell hybrids containing defined sets of
5   human chromosomes were analyzed by Southern blotting and hybridization with the VEGF-C cDNA probe. Among 24 DNA samples on the hybrid panel, representing different human chromosomes, human-specific signals were observed only in hybrids which contained human chromosome 4. The results
10 were confirmed by PCR of somatic cell hybrid DNAs using VEGF-C specific primers, where amplified bands were obtained only from DNAs containing human chromosome 4.

A genomic P1 plasmid for VEGF-C was isolated using specific primers and PCR and verified by Southern
15 blotting and hybridization using a VEGF-C specific cDNA probe. The chromosomal localization of VEGF-C was further studied using metaphase FISH. Using the P1 probe for VEGF-C in FISH a specific hybridization to the 4q34 chromosomal band was detected in 40 out of 44 metaphases.
20 Double-fluorochrome hybridization using a cosmid probe specific for the aspartylglucosaminidase (AGA) gene showed that VEGF-C is located just proximal to the AGA gene previously mapped to the 4q34-35 chromosomal band.

Biotin labelled VEGF-C P1 and digoxigenin
25 labeled AGA cosmid probes were hybridized simultaneously to metaphase chromosomes. This experiment demonstrated that the AGA gene is more telomerically located than the VEGF-C gene. The foregoing example demonstrates the utility of polynucleotides of the invention as
30 chromosomal markers and for the presence or absence of the VEGF-C gene region in normal or diseased cells. The VEGF-C locus at 4q34 is a candidate target for mutations leading to vascular malformations or cardiovascular diseases.

- 63 -

EXAMPLE 18

Effect of glucose concentration and hypoxia on VEGF, VEGF-B and VEGF-C mRNA levels in C6 glioblastoma cells

5   Confluent cultures of C6 cells (ATCC CCL 107) were grown on 10 cm diameter tissue culture plates containing 2.5 ml of DMEM and 5% fetal calf serum plus antibiotics. The cultures were exposed for 16 hours to normoxia in a normal cell culture incubator containing
10   5% CO_2 or hypoxia by closing the culture plates in an airtight glass chamber and burning a piece of wood inside until the flame was extinguished due to lack of oxygen. Polyadenylated RNA was isolated (as in the other examples), and 8 micrograms of the RNA was
15   electrophoresed and blot-hybridized with a mixture of the VEGF, VEGF-B and VEGF-C probes (see Fig. 12). The results show that hypoxia strongly induces VEGF mRNA expression, both in low and high glucose, but has no significant effect on the VEGF-B mRNA levels. The VEGF-C
20   mRNA isolated from hypoxic cells runs slightly faster in gel electrophoresis and an extra band of faster mobility can be seen below the upper mRNA band. This observation suggests that hypoxia affects VEGF-C RNA processing. One explanation for this observation is that VEGF-C mRNA
25   splicing is altered, affecting the VEGF-C open reading frame and resulting in an alternative VEGF-C protein being produced by hypoxic cells. Such alternative forms of VEGF-C and VEGF-C-encoding polynucleotides are contemplated as an aspect of the invention. This data
30   indicates screening and diagnostic utilities for polynucleotides and polypeptides of the invention, such as methods whereby a biological sample is screened for the hypoxia-induced form of VEGF-C and/or VEGF-C mRNA. The data further suggests a therapeutic indication for
35   antibodies and/or other inhibitors of the hypoxia-induced form of VEGF-C or the normal form of VEGF-C.

- 64 -

EXAMPLE 19

Pulse-chase labeling and immunoprecipitation of VEGF-C polypeptides from 293 EBNA cells transfected with VEGF-C expression vector.

5   The following VEGF-C branched amino-terminal peptide, designated PAM126, was synthesized for production of anti-VEGF-C antiserum:

NH$_2$-E-E-T-I-K-F-A-A-A-H-Y-N-T-E-I-L-K-COOH (SEQ ID NO: 39).

10   In particular, PAM126 was synthesized as a branched polylysine structure K3PA4 having four peptide acid (PA) chains attached to two available lysine (K) residues. The synthesis was performed on a 433A Peptide Synthesizer (Applied Biosystems) using Fmoc-chemistry and TentaGel S

15   MAP RAM10 resin mix (RAPP Polymere GmbH, Tubingen, Germany), yielding both cleavable and resin-bound peptides. The cleavable peptide was purified via reverse phase HPLC and was used together with the resin-bound peptide in immunizations. The correctness of the

20   synthesis products were confirmed using mass-spectroscopy (Lasermatt).

          The PAM126 peptide was dissolved in phosphate buffered saline (PBS), mixed with Freund's adjuvant, and used for immunization of rabbits at bi-weekly intervals

25   using methods standard in the art (Harlow and Lane, *Antibodies, a laboratory manual*, Cold Spring Harbor Laboratory Press (1988)). Antisera obtained after the fourth booster immunization was used for immunoprecipitation of VEGF-C in pulse-chase experiments,

30   as described below.

          For pulse-chase analysis, 293 EBNA cells transfected with a VEGF-C expression vector (i.e., the FLT4-L cDNA inserted into the pREP7 expression vector as described above) were incubated for 30 minutes in

35   methionine-free, cysteine-free, serum-free DMEM culture medium at 37°C. The medium was then changed, and 200 μCi of Pro-mix™ (Amersham), was added. The cell layers were incubated in this labeling medium for two hours, washed

with PBS, and incubated for 0, 15, 30, 60, 90, 120, or
180 minutes in serum-free DMEM (chase). After the
various chase periods, the medium was collected, the
cells were again washed two times in PBS, and lysed in
5   immunoprecipitation buffer. The VEGF-C polypeptides were
analyzed from both the culture medium and from the cell
lysates by immunoprecipitation, using the VEGF-C-specific
antiserum raised against the NH$_2$-terminal peptide (PAM126)
of the 23 kD VEGF-C form. Immunoprecipitated
10   polypeptides were analyzed via SDS-PAGE followed by
autoradiography.

Referring to Fig. 19, the resultant
autoradiograms demonstrate that immediately after a 2
hour labeling (chase time 0), the VEGF-C vector-
15   transfected cells contained a radioactive 55 kD
polypeptide band, which is not seen in mock-transfected
cells (M). This 55 kD polypeptide band gradually
diminishes in intensity with increasing chase periods,
and is no longer detected in the cells by 180 minutes of
20   chase. A 32 kD polypeptide band also is observed in
VEGF-C transfected cells (and not mock-transfected
cells). This 32 kD band disappears with similar kinetics
to that of the 55 kD band. Simultaneously, increasing
amounts of 32 kD (arrow) and subsequently 23 kD (arrow)
25   and 14 kD polypeptides appear in the medium.

Collectively, the data from the pulse-chase
experiments indicate that the 55 kD intracellular
polypeptide represents a pro-VEGF-C polypeptide, which is
not secreted from cells, but rather is first
30   proteolytically cleaved into the 32 kD form. The 32 kD
form is secreted and simultaneously further processed by
proteolysis into the 23 kD and 14 kD forms. Without
intending to be limited to a particular theory, it is
believed that processing of the VEGF-C precursor occurs
35   as removal of a signal sequence, removal of the
COOH-terminal domain (BR3P), and removal of an amino
terminal polypeptide, resulting in a VEGF-C polypeptide

having the TEE... amino terminus.

At high resolution, the 23 kD polypeptide band
appears as a closely-spaced polypeptide doublet,
suggesting heterogeneity in cleavage or glycosylation.

EXAMPLE 20
Isolation of Mouse and Quail
cDNA Clones Encoding VEGF-C

To clone a mouse variant of VEGF-C,
approximately 1 x 10^6 bacteriophage lambda clones of a
commercially-available 12 day mouse embryonal cDNA
library (lambda EXlox library, Novagen, catalog number
69632-1) were screened with a radiolabeled fragment of
human VEGF-C cDNA containing nucleotides 495 to 1661 of
SEQ ID NO: 32. One positive clone was isolated.

A 1323 bp EcoRI/HindIII fragment of the insert
of the isolated mouse cDNA clone was subcloned into the
corresponding sites of the pBluescript SK+ vector
(Stratagene) and sequenced. The cDNA sequence of this
clone was homologous to the human VEGF-C sequence
reported herein, except that about 710 bp of 5'-end
sequence present in the human clone was not present in
the mouse clone.

For further screening of mouse cDNA libraries,
a HindIII-BstXI (HindIII site is from the pBluescript SK+
polylinker) fragment of 881 bp from the coding region of
the mouse cDNA clone was radiolabeled and used as a probe
to screen two additional mouse cDNA libraries. Two
additional cDNA clones from an adult mouse heart ZAP II
cDNA library (Stratagene, catalog number 936306) were
identified. Three additional clones also were isolated
from a mouse heart 5'-stretch-plus cDNA library in λgt11
(Clontech Laboratories, Inc., catalog number ML5002b).
Of the latter three clones, one was found to contain an
insert of about 1.9 kb. The insert of this cDNA clone
was subcloned into EcoRI sites of pBluescript SK+ vector
and both strands of this clone were completely sequenced,
resulting in the nucleotide and deduced amino acid

- 67 -

sequences shown in SEQ ID NOs: 40 and 41.

It is contemplated that the polypeptide corresponding to SEQ ID NO: 41 is processed into a mature mouse VEGF-C protein, in a manner analogous to the

5   processing of the human VEGF-C prepropeptide. Putative cleavage sites for the mouse protein are identified using procedures outlined above for identification of cleavage sites for the human VEGF-C polypeptide.

The foregoing results demonstrate the utility

10   of polynucleotides of the invention for identifying and isolating polynucleotides encoding other non-human mammalian variants of VEGF-C. Such identified and isolated polynucleotides, in turn, can be expressed (using procedures similar to those described in preceding

15   examples) to produce recombinant polypeptides corresponding to non-human mammalian variants of VEGF-C.

The mouse and human VEGF-C sequences were used to design probes for isolating a quail VEGF-C cDNA from a quail cDNA library. A fragment of the human VEGF-C cDNA

20   comprising nucleotides 495-1670 of SEQ ID NO: 32 was obtained by PCR amplification, cloned into the pCRII vector (Invitrogen) according to the manufacturer's instructions, and amplified. The insert was isolated by *Eco* RI digestion and preparative gel electrophoresis and

25   then labelled using radioactive dCTP and random priming. A cDNA library made from quail embryos of stage E-4 in pcDNA-1 vector (Invitrogen) was then screened using this probe. About 200,000 colonies were plated and filter replicas were hybridized with the radioactive probe.

30   Nine positive clones were identified and secondarily plated. Two of the nine clones hybridized in secondary screening. The purified clones (clones 1 and 14) had approximately 2.7 kb *Eco* RI inserts. Both clones were amplified and then sequenced using the T7 and SP6 primers

35   (annealing to the vector). In addition, an internal *Sph* I restriction endonuclease cleavage site was identified

- 68 -

about 1.9 kb from the T7 primer side of the vector and
used for subcloning 5'- and 3'- Sph I fragments, followed
by sequencing from the Sph I end of the subclones. The
sequences obtained were identical from both clones and

5   showed a high degree of similarity to the human VEGF-C
coding region. Subsequently, walking primers were made
in both directions and double-stranded sequencing was
completed for 1743 base pairs, including the full-length
open reading frame.

10   The cDNA sequence obtained includes a long open
reading frame and 5' untranslated region. The DNA and
deduced amino acid sequences for the quail cDNA are set
forth in SEQ ID NOs: 52 and 53, respectively. As shown
in Fig. 8, the human, murine, and avian (quail) VEGF-C

15   precursor amino acid sequences share a significant degree
of conservation. This high degree of homology permits
the isolation of VEGF-C encoding sequences from other
species, especially vertebrate species, and more
particularly mammalian and avian species, using

20   polynucleotides of the present invention as probes and
using standard molecular biological techniques such as
those described herein.

                        EXAMPLE 21
                N-terminal peptide sequence analyses
25   of recombinant VEGF-C
              Cells (293 EBNA) transfected with VEGF-C cDNA
      (see Example 13) secrete several forms of recombinant
      VEGF-C (Fig. 21A, lane IP). In the absence of
      alkylation, the three major, proteolytically-processed

30   forms of VEGF-C migrate in SDS-PAGE as proteins with
      apparent molecular masses of 32/29 kD (doublet), 21 kD
      and 15 kD. Two minor polypeptides exhibit approximate
      molecular masses of 63 and 52 kD, respectively. One of
      these polypeptides is presumably a glycosylated and non-

35   processed form; the other polypeptide is presumably
      glycosylated and partially processed.
              To determine sites of proteolytic cleavage of

the VEGF-C precursor, an immunoaffinity column was used
to purify VEGF-C polypeptides from the conditioned medium
of 293 EBNA cells transfected with VEGF-C cDNA. To
prepare the immunoaffinity column, a rabbit was immunized
5　　with a synthetic peptide corresponding to amino acids
104-120 of SEQ ID NO: 33: H_2N-EETIKFAAAHYNTEILK (see
PAM126 in Example 19). The IgG fraction was isolated
from the serum of the immunized rabbit using protein A
Sepharose (Pharmacia). The isolated IgG fraction was
10　　covalently bound to CNBr-activated Sepharose CL-4B
(Pharmacia) using standard techniques at a concentration
of 5 mg IgG/ml of Sepharose. This immunoaffinity matrix
was used to isolate processed VEGF-C from 1.2 liters of
the conditioned medium (CM).

15　　　　　The purified material eluted from the column
was analyzed by gel electrophoresis and Western blotting.
Fractions containing VEGF-C polypeptides were combined,
dialyzed against 10 mM Tris HCl, vacuum-dried,
electrotransferred to Immobilon-P (polyvinylidene
20　　difluoride or PVDF) transfer membrane (Millipore,
Marlborough, MA) and subjected to N-terminal amino acid
sequence analysis.

　　　　　The polypeptide band of 32 kD yielded two
distinct sequences: NH_2-FESGLDLSDA... and NH_2-
25　　AVVMTQTPAS... (SEQ ID NO: 51), the former corresponding
to the N-terminal part of VEGF-C after cleavage of the
signal peptide, starting from amino acid 32 (SEQ ID NO:
33), and the latter corresponding to the kappa-chain of
IgG, which was present in the purified material due to
30　　"leakage" of the affinity matrix during the elution
procedure.

　　　　　In order to obtain the N-terminal peptide
sequence of the 29 kD form of VEGF-C, a construct (VEGF-C
NHis) encoding a VEGF-C variant was generated. In
35　　particular, the construct encoded a VEGF-C variant that
fused a 6xHis tag to the N-terminus of the secreted
precursor (i.e., between amino acids 31 and 33 in SEQ ID

- 70 -

NO: 33). The phenylalanine at position 32 was removed to
prevent possible cleavage of the tag sequence during
secretion of VEGF-C. The VEGF-C NHis construct was
cloned into pREP7 as a vector; the construction is
5   described more fully in Example 28, below.

The calcium phosphate co-precipitation
technique was used to transfect VEGF-C NHis into 293 EBNA
cells. Cells were incubated in DMEM/10% fetal calf serum
in 15 cm cell culture dishes (a total of 25 plates). The
10   following day, the cells were reseeded into fresh culture
dishes (75 plates) containing the same medium and
incubated for 48 hours. Cell layers were then washed
once with PBS and DMEM medium lacking FCS was added.
Cells were incubated in this medium for 48 hours and the
15   medium was collected, cleared by centrifugation at 5000 x
g and concentrated 500X using an Ultrasette Tangential
Flow Device (Filtron, Northborough, MA), as described in
Example 5 above. VEGF-C NHis was purified from the
concentrated conditioned medium using TALON™ Metal
20   Affinity Resin (Clontech Laboratories, Inc.) and the
manufacturer's protocol for native protein purification
using imidazole-containing buffers. The protein was
eluted with a solution containing 20 mM Tris-HCl (pH
8.0), 100 mM NaCl, and 200 mM imidazole. The eluted
25   fractions containing purified VEGF-C NHis were detected
by immunoblotting with Antiserum 882 (antiserum from
rabbit 882, immunized with the PAM-126 polypeptide).
Fractions containing VEGF-C NHis were combined, dialyzed
and vacuum-dried. As can be seen in Fig. 27, due to the
30   presence of the 6xHis tag at the N-terminus of this form
of VEGF-C, the upper component of the major doublet of
the VEGF-CNHis migrates slightly slower than the 32 kD
form of wild type VEGF-C, thereby improving the
separation of the VEGF-CNHis 32 kD variant from the 29 kD
35   band using SDS-PAGE. Approximately 15 μg of the purified
VEGF-C were subjected to SDS-PAGE under reducing
conditions, electrotransferred to Immobilon-P (PVDF)

- 71 -

transfer membrane (Millipore, Inc., Marlborough, MA) and
the band at 29 kD was subjected to N-terminal amino acid
sequence analysis. This sequence analysis revealed an N-
terminal sequence of H_2N-SLPAT . . ., corresponding to
amino acids 228-232 of VEGF-C (SEQ ID NO: 33).

The polypeptide band of 21 kD yielded the sequence
H_2N-AHYNTEILKS . . ., corresponding to an amino-terminus
starting at amino acid 112 of SEQ ID NO: 33. Thus, the
proteolytic processing site which results in the 21 kD
form of VEGF-C produced by transfected 293 EBNA cells
apparently occurs nine amino acid residues downstream of
the cleavage site which results in the 23 kD form of
VEGF-C secreted by PC-3 cells.

The N-terminus of the 15 kD form was identical
to the N-terminus of the 32 kD form (NH_2-FESGLDLSDA...).
The 15 kD form was not detected when recombinant VEGF-C
was produced by COS cells. This suggests that production
of this form is cell lineage specific.

Example 22
Dimeric and monomeric forms of VEGF-C

The composition of VEGF-C dimers was analyzed
as follows. Cells (293 EBNA cells), transfected with the
pREP7 VEGF-C vector as described in Example 11, were
metabolically labelled with Pro-mix L-[^{35}S] labelling mix
(Amersham Corp.) to a final concentration of 100 μCi/ml.

In parallel, a VEGF-C mutant, designated
"R102S", was prepared and analyzed. To prepare the DNA,
encoding VEGF-C-R102S, the arginine codon at position 102
of SEQ ID NO: 33 was replaced with a serine codon. This
VEGF-C-R102S-encoding DNA, in a pREP7 vector, was
transfected into 293 EBNA cells and expressed as
described above. VEGF-C polypeptides were
immunoprecipitated using antisera 882 (obtained by
immunization of a rabbit with a polypeptide corresponding
to residues 104-120 of SEQ ID NO: 33 (see previous
Example)) and antisera 905 (obtained by immunization of a

rabbit with a polypeptide corresponding to a portion of
the prepro- VEGF-C leader: H₂N-ESGLDLSDAEPDAGEATAYASK
(residues 33 to 54 of SEQ ID NO: 33).

5   The immunoprecipitates from each cell culture
were subjected to SDS-PAGE under non-denaturing
conditions (Fig. 21B). Bands 1-6 were cut out from the
gel, soaked for 30 minutes in 1x gel-loading buffer
containing 200 mM β-mercaptoethanol, and individually
subjected to SDS-PAGE under denaturing conditions (Figs.
10   21A and 21C, lanes 1-6).

         As can be seen from Figures 21A-C, each high
molecular weight form of VEGF-C (Fig. 21B, bands 1-4)
consists of at least two monomers bound by disulfide
bonds (Compare Figs. 21A and 21C, lanes 1-4, in the
15   reducing gels). The main component of bands 1-3 is the
doublet of 32/29 kD, where both proteins are present in
an equimolar ratio. The main fraction of the 21 kD form
is secreted as either a monomer or as a homodimer
connected by means other than disulfide bonds (bands 6
20   and lanes 6 in Figs. 21A-C).

         The R102S mutation creates an additional site
for N-linked glycosylation in VEGF-C at the asparagine
residue at position 100 in SEQ ID NO: 33. Glycosylation
at this additional glycosylation site increases the
25   apparent molecular weight of polypeptides containing the
site, as confirmed in Figures 21A-C and Figures 22A-B.
The additional glycosylation lowers the mobility of forms
of VEGF-C-R102S that contain the additional glycosylation
site, when compared to polypeptides of similar primary
30   structure corresponding to VEGF-C. Figures 21A-C and
Figures 22A-B reveal that the VEGF-C-R102S polypeptides
corresponding to the 32 kD and 15 kD forms of wt VEGF-C
exhibit increased apparent molecular weights, indicating
that each of these polypeptides contains the newly
35   introduced glycosylation site. In particular, the VEGF-
C-R102S polypeptide corresponding to the 15 kD
polypeptide from VEGF-C comigrates on a gel with the 21

kD form of the wild type (wt) VEGF-C, reflecting a shift
on the gel to a position corresponding to a greater
apparent molecular weight. (Compare lanes 4 in Figures
21A and 21C).

5   In a related experiment, another VEGF-C mutant,
designated "R226,227S," was prepared and analyzed. To
prepare a DNA encoding VEGF-C-R226,227S, the arginine
codons at positions 226 and 227 of SEQ ID NO: 33 were
replaced with serine codons by site-directed mutagenesis.
10 The resultant DNA was transfected into 293 EBNA cells as
described above and expressed and analyzed under the same
conditions as described for VEGF-C and VEGF-C-R102S. In
the conditioned medium from the cells expressing VEGF-C-
R226,227S, no 32 kD form of VEGF-C was detected. These
15 results indicate that a C-terminal cleavage site of wild-
type VEGF-C is adjacent to residues 226 and 227 of SEQ ID
NO: 33, and is destroyed by the mutation of the arginines
to serines. Again, the mobility of the 29 kD component
of the doublet was unchanged (Figures 22A-B).

20   Taken together, these data indicate that the
major form of the processed VEGF-C is a heterodimer
consisting of (1) a polypeptide of 32 kD containing
amino acids 32-227 of the prepro-VEGF-C (amino acids 32
to 227 in SEQ ID NO: 33) attached by disulfide bonds to
25 (2) a polypeptide of 29 kD beginning with amino acid 228
in SEQ ID NO: 33. These data are also supported by a
comparison of the pattern of immunoprecipitated, labelled
VEGF-C forms using antisera 882 and antisera 905.

          When VEGF-C immunoprecipitation was carried out
30 using conditioned medium, both antisera (882 and 905)
recognized some or all of the three major processed forms
of VEGF-C (32/29 kD, 21 kD and 15 kD). When the
conditioned medium was reduced by incubation in the
presence of 10 mM dithiothreitol for two hours at room
35 temperature with subsequent alkylation by additional
incubation with 25 mM iodoacetamide for 20 minutes at
room temperature, neither antibody precipitated the 29 kD

component, although antibody 882 still recognized
polypeptides of 32 kD, 21 kD and 15 kD. These results
are consistent with the nature of the oligopeptide
antigen used to elicit the antibodies contained in
antisera 882, an oligopeptide containing amino acid
residues 104-120 of SEQ ID NO: 33. On the other hand,
antisera 905 recognized only the 32 kD and 15 kD
polypeptides, which include sequence of the oligopeptide
(amino acids 33 to 54 of SEQ ID NO: 33) used for
immunization to obtain antisera 905. Taking into account
the mobility shift of the 32 kD and 15 kD forms, the
immunoprecipitation results with the R102S mutant were
similar (Figs. 23A-B). The specificity of antibody 905
is confirmed by the fact that it did not recognize a
VEGF-C ΔN variant form wherein the N-terminal propeptide
spanning residues 32-102 of the unprocessed polypeptide
had been deleted (Fig. 23B).

The results of these experiments also
demonstrate that the 21 kD polypeptide is found (1) in
heterodimers with other molecular forms (see Figs. 21A-C
and Figs. 22A-B), and (2) secreted as a monomer or a
homodimer held by bonds other than disulfide bonds (Figs.
21A and 21B, lanes 6).

The experiments disclosed in this example
demonstrate that several forms of VEGF-C exist. A
variety of VEGF-C monomers were observed and these
monomers can vary depending on the level and pattern of
glycosylation. In addition, VEGF-C was observed as a
multimer, for example a homodimer or a heterodimer. The
processing of VEGF-C is schematically presented in Fig.
18 (disulfide bonds not shown). All forms of VEGF-C are
within the scope of the present invention.

Example 23
In situ Hybridization of Mouse Embryos
To analyze VEGF-C mRNA distribution in
different cells and tissues, sections of 12.5 and 14.5-

- 75 -

day post-coitus (p.c.) mouse embryos were prepared and
analyzed via *in situ* hybridization using labeled VEGF-C
probes. *In situ* hybridization of tissue sections was
performed as described in Västrik *et al., J. Cell Biol.,*
5   *128*:1197-1208 (1995). A mouse VEGF-C antisense RNA probe
was generated from linearized pBluescript II SK+ plasmid
(Stratagene Inc., La Jolla, CA), containing a cDNA
fragment corresponding to nucleotides 499-979 of a mouse
VEGF-C cDNA (SEQ ID NO: 40). Radiolabeled RNA was
10   synthesized using T7 polymerase and [^{35}S]-UTP (Amersham).
Mouse VEGF-B antisense and sense RNA probes were
synthesized in a similar manner from linearized pCRII
plasmid containing the mouse VEGF-B cDNA insert as
described Olofsson *et al., Proc. Natl. Acad. Sci. (USA),*
15   *93*:2576-2581 (1996). The high stringency wash was for 45
minutes at 65°C in a solution containing 30 mM
dithiothreitol (DTT) and 4 x SSC. The slides were
exposed for 28 days, developed and stained with
hematoxylin. For comparison, similar sections were
20   hybridized with a VEGFR-3 probe and the 12.5-day p.c.
embryos were also probed for VEGF-B mRNA.

          Figures 34A-D show darkfield (Figures 34A-C)
and lightfield (Figure 34D) photomicrographs of 12.5 day
p.c. embryo sections probed with the antisense (Fig. 34A)
25   and sense (Figs. 34C-D) VEGF-C probes. Fig. 34A
illustrates a parasagittal section, where VEGF-C mRNA is
particularly prominent in the mesenchyme around the
vessels surrounding the developing metanephros (mn). In
addition, hybridization signals were observed between the
30   developing vertebrae (vc), in the developing lung
mesenchyme (lu), in the neck region and developing
forehead. The specificity of these signals is evident
from the comparison with VEGF-B expression in an adjacent
section (Fig. 34B), where the myocardium gives a very
35   strong signal and lower levels of VEGF-B mRNA are
detected in several other tissues. Both genes appear to
be expressed in between the developing vertebrae (vc), in

- 76 -

the developing lung (lu) and forehead. Hybridization of
the VEGF-C sense probe showed no specific expression
within these structures (Fig. 34C).

5   Figs. 35A-D show a comparison of the expression
patterns of VEGF-C and VEGFR-3 in 12.5 day p.c. mouse
embryos in the jugular region, where the developing
dorsal aorta and cardinal vein are located. This is the
area where the first lymphatic vessels sprout from venous
sac-like structures according to the long-standing theory
10   of Sabin, *Am. J. Anat.*, *9*:43-91 (1909). As can be seen
from Figs. 35A-D, an intense VEGF-C signal is detected in
the mesenchyme surrounding the developing venous sacs
(Figs. 35A and 35C) which are positive for VEGFR-3 (Figs.
35B and 35D).

15   The mesenterium supplies the developing gut
with blood and contains developing lymphatic vessels.
The developing 14.5 day p.c. mesenterium is positive for
VEGF-C mRNA, with particularly high expression in
connective tissue surrounding certain vessels (arrowheads
20   in Figs. 35E-H). This signal in Fig. 35E should be
distinguished from the false positive reflection of light
from red blood cells within the vessel. The adjacent
mesenterial VEGFR-3 signals shown in Fig. 35F originate
from small capillaries of the mesenterium (arrowhead).
25   Therefore, there appears to be a paracrine relationship
between the production of the mRNAs for VEGF-C and its
receptor. This data indicates that VEGF-C is expressed
in a variety of tissues. Moreover, the pattern of
expression is consistent with a role for VEGF-C in venous
30   and lymphatic vessel development. Further, the data
reveals that VEGF-C is expressed in non-human animals.

Example 24
Analysis of VEGF, VEGF-B, and VEGF-C mRNA Expression
in Fetal and Adult Tissues

35   A human fetal tissue Northern blot containing 2
μg of polyadenylated RNAs from brain, lung, liver and
kidney (Clontech Inc.) was hybridized with a pool of the

- 77 -

following probes: a human full-length VEGF-C cDNA insert
(Genbank Acc. No. X94216), a human VEGF-B$_{167}$ cDNA fragment
(nucleotides 1-382, Genbank Acc. No. U48800) obtained by
PCR amplification; and a human VEGF 581 bp cDNA fragment
5   covering base pairs 57-638 (Genbank Acc. No. X15997).
Blots were washed under stringent conditions, using
techniques standard in the art.

        Mouse embryo multiple tissue Northern blot
(Clontech Inc.) containing 2 µg of polyadenylated RNAs
10   from 7, 11, 15 and 17 day postcoital (p.c.) embryos was
hybridized with mouse VEGF-C cDNA fragment (base pairs
499-656). A mouse adult tissue Northern blot was
hybridized with the probes for human VEGF, VEGF-B$_{167}$,
VEGF-C and with a VEGFR-3 cDNA fragment (nucleotides 1-
15   595; Genbank Acc. No. X68203).

        In adult mouse tissues, both 2.4 kb and 2.0 kb
mRNA signals were observed with the VEGF-C probe, at an
approximately 4:1 ratio. The most conspicuous signals
were obtained from lung and heart RNA, while kidney,
20   liver, brain, and skeletal muscle had lower levels, and
spleen and testis had barely visible levels. As in the
human tissues, VEGF mRNA expression in adult mice was
most abundant in lung and heart RNA, whereas the other
samples showed less coordinate regulation with VEGF-C
25   expression. Skeletal muscle and heart tissues gave the
highest VEGF-B mRNA levels from adult mice, as previously
reported Olofsson et al., Proc. Natl. Acad. Sci. (USA),
93:2576-2581 (1996). Comparison with VEGFR-3 expression
showed that the tissues where VEGF-C is expressed also
30   contain mRNA for its cognate receptor tyrosine kinase,
although in the adult liver VEGFR-3 mRNA was
disproportionally abundant.

        To provide a better insight into the regulation
of the VEGF-C mRNA during embryonic development,
35   polyadenylated RNA isolated from mouse embryos of various
gestational ages (7, 11, 15, and 17 day p.c.) was
hybridized with the mouse VEGF-C probe. These analyses

- 78 -

showed that the amount of 2.4 kb VEGF-C mRNA is
relatively constant throughout the gestational period.

Example 25

Regulation of mRNAs for VEGF family members by serum, interleukin-1 and dexamethasone in human fibroblasts in culture

Human IMR-90 fibroblasts were grown in DMEM
medium containing 10% FCS and antibiotics. The cells
were grown to 80% confluence, then starved for 48 hours
in 0.5 % FCS in DMEM. Thereafter, the growth medium was
changed to DMEM containing 5% FCS, with or without 10
ng/ml interleukin-1 (IL-1) and with or without 1 mM
dexamethasone, as indicated in Figs. 24A-B. The culture
plates were incubated with these additions for the times
indicated, and total cellular RNA was isolated using the
TRIZOL kit (GIBCO-BRL). About 20 μg of total RNA from
each sample was electrophoresed in 1.5% formaldehyde-
agarose gels as described in Sambrook et al., supra
(1989). The gel was used for Northern blotting and
hybridization with radiolabeled insert DNA from the human
VEGF clone (a 581 bp cDNA covering bps 57-638, Genbank
Acc. No. 15997) and a human VEGF-B$_{167}$ cDNA fragment
(nucleotides 1-382, Genbank Acc. No. U48800) (Fig. 25B).
Subsequently, the Northern blots were probed with
radiolabelled insert from the VEGF-C cDNA plasmid (Fig.
24A). Primers were labelled using a standard technique
involving enzymatic extension reactions of random
primers, as would be understood by one of ordinary skill
in the art. The mobilities of the 28S and 18S ribosomal
RNA bands are indicated, based on UV photography of
ethidium bromide stained RNA before the transfer.

As can be seen in Figs. 24A-B, very low levels
of VEGF-C and VEGF are expressed by the starved IMR-90
cells as well as cells after 1 hour of stimulation. In
contrast, abundant VEGF-B mRNA signal is visible under
these conditions. After a 4 hours of serum stimulation,
there is a strong induction of VEGF-C and VEGF mRNAs,

which are further increased in the IL-1 treated sample.
The effect of IL-1 seems to be abolished in the presence
of dexamethasone. A similar pattern of enhancement is
maintained in the 8 hour sample, but a gradual down-
regulation of all signals occurs for both RNAs in the 24
hour and 48 hour samples. In contrast, VEGF-B mRNA
levels remain constant and thus show remarkable stability
throughout the time period. The results are useful in
guiding efforts to use VEGF-C and its fragments, its
antagonists, and anti-VEGF-C antibodies in methods for
treating a variety of disorders.

Example 26
Expression and analysis of recombinant murine VEGF-C

The mouse VEGF-C cDNA was expressed as a
recombinant protein and the secreted protein was analyzed
for its receptor binding properties. The binding of
mouse VEGF-C to the human VEGFR-3 extracellular domain
was studied by using media from Bosc23 cells transfected
with mouse VEGF-C cDNA in a retroviral expression vector.

The 1.8 kb mouse VEGF-C cDNA was cloned as an
EcoRI fragment into the retroviral expression vector
pBabe-puro containing the SV40 early promoter region
[Morgenstern et al., Nucl. Acids Res., 18:3587-3595
(1990)], and transfected into the Bosc23 packaging cell
line [Pearet et al., Proc. Natl. Acad. Sci. (USA),
90:8392-8396 (1994)] by the calcium-phosphate
precipitation method. For comparison, Bosc23 cells also
were transfected with the previously- described human
VEGF-C construct in the pREP7 expression vector. The
transfected cells were cultured for 48 hours prior to
metabolic labelling. Cells were changed into DMEM medium
devoid of cysteine and methionine, and, after 45 minutes
of preincubation and medium change, Pro-mix™ L-[^{35}S] in
vitro cell labelling mix (Amersham Corp.), in the same
medium, was added to a final concentration of about 120
μCi/ml. After 6 hours of incubation, the culture medium

- 80 -

was collected and clarified by centrifugation.

For immunoprecipitation, 1 ml aliquots of the
media from metabolically-labelled Bosc23 cells
transfected with empty vector or mouse or human
5   recombinant VEGF-C, respectively, were incubated
overnight on ice with 2 μl of rabbit polyclonal antiserum
raised against an N-terminal 17 amino acid oligopeptide
of mature human VEGF-C (H_2N-EETIKFAAAHYNTEILK) (SEQ ID NO:
33, residues 104-120). Thereafter, the samples were
10   incubated with protein A sepharose for 40 minutes at 4°C
with gentle agitation. The sepharose beads were then
washed twice with immunoprecipitation buffer and four
times with 20 mM Tris-HCl, pH 7.4. Samples were boiled
in Laemmli buffer and analyzed by 12.5% sodium dodecyl
15   sulfate polyacrylamide gel electrophoresis (SDS-PAGE).

Immunoprecipitation of VEGF-C from media of
transfected and metabolically-labelled cells revealed
bands of approximately 30-32x10³ M_r (a doublet) and 22-
23x10³ M_r in 12.5% SDS-PAGE. These bands were not
20   detected in samples from nontransfected or mock-
transfected cells as shown in Fig. 32 (*i.e.,* lanes marked
"vector"). These results show that antibodies raised
against human VEGF-C recognize the corresponding mouse
ligand.

25   For receptor binding experiments, 1 ml aliquots
of media from metabolically-labelled Bosc23 cells were
incubated with VEGFR-3 extracellular domain (see Example
3), covalently coupled to sepharose, for 4 hours at 4°C
with gentle mixing. The sepharose beads were washed
30   four times with ice-cold phosphate buffered saline (PBS),
and the samples were analyzed by gel electrophoresis as
described in Joukov et al., *EMBO J.,* *15:*290-298 (1996).

As can be seen from Fig. 32, similar 30-32 x 10³
M_r doublet and 22-23 x 10³ M_r polypeptide bands were
35   obtained in the receptor binding assay as compared to the
immunoprecipitation assay. Thus, mouse VEGF-C binds to
human VEGFR-3. The slightly faster mobility of the mouse

VEGF-C polypeptides may be caused by the four amino acid
residue difference observed in sequence analysis
(residues H88-E91, Fig. 31).

      The capacity of mouse recombinant VEGF-C to
5   induce VEGFR-3 autophosphorylation was also investigated.
For the VEGFR-3 receptor stimulation experiments,
subconfluent NIH 3T3-Flt4 cells, Pajusola et al.,
Oncogene, 9:3545-3555 (1994), were starved overnight in
serum-free medium containing 0.2% BSA. In general, the
10   cells were stimulated with the conditioned medium from
VEGF-C vector-transfected cells for 5 minutes, washed
three times with cold PBS containing 200 μM vanadate, and
lysed in RIPA buffer for immunoprecipitation analysis.
The lysates were centrifuged for 25 minutes at 16000 x g
15   and the resulting supernatants were incubated for 2 hours
on ice with the specific antisera, followed by
immunoprecipitation using protein A-sepharose and
analysis in 7% SDS-PAGE. Polypeptides were transferred
to nitrocellulose and analyzed by immunoblotting using
20   anti-phosphotyrosine (Transduction Laboratories) and
anti-receptor antibodies, as described by Pajusola et
al., Oncogene, 9:3545-3555 (1994). Filter stripping was
carried out at 50°C for 30 minutes in 100 mM 2-
mercaptoethanol, 2% SDS, 62.5 mM Tris-HCl, pH 6.7, with
25   occasional agitation. The results of the experiment are
shown in Fig. 33. The results demonstrate that culture
medium containing mouse VEGF-C stimulates the
autophosphorylation of VEGFR-3 to a similar extent as
human baculoviral VEGF-C or the tyrosyl phosphatase
30   inhibitor pervanadate.

      VEGFR-2 stimulation was studied in subconfluent
porcine aortic endothelial (PAE) cells expressing Kdr
(VEGFR-2) (PAE-VEGFR-2) [Waltenberger et al., J. Biol.
Chem., 269:26988-26995 (1994)], which were starved
35   overnight in serum-free medium containing 0.2% BSA.
Stimulation was carried out and the lysates prepared as
described above. For receptor immunoprecipitation,

- 82 -

specific antiserum for VEGFR-2 [Waltenberger et al., J.
Biol. Chem., 269:26988-26995 (1994)] was used. The
immunoprecipitates were analyzed as described for VEGFR-3
in 7% SDS-PAGE followed by Western blotting with anti-
5   phosphotyrosine antibodies, stripping of the filter, and
re-probing it with anti-VEGFR-2 antibodies (Santa Cruz).
        Mouse VEGF-C appeared to be a potent inducer of
VEGFR-3 autophosphorylation, with the $195x10^3$ M_r precursor
and proteolytically cleaved $125x10^3$ M_r tyrosine kinase
10   polypeptides of the receptor (Pajusola et al., Oncogene,
9:3545-3555 (1994)), being phosphorylated. VEGFR-2
stimulation was first tried with unconcentrated medium
from cells expressing recombinant VEGF-C, but
immunoblotting analysis did not reveal any receptor
15   autophosphorylation.
        To further determine whether mouse recombinant
VEGF-C can also induce VEGFR-2 autophosphorylation as
observed for human VEGF-C, PAE cells expressing VEGFR-2
were stimulated with tenfold concentrated medium from
20   cultures transfected with mouse VEGF-C expression vector
and autophosphorylation was analyzed. For comparison,
cells treated with tenfold concentrated medium containing
human recombinant VEGF-C (Joukov et al., (1996)),
unconcentrated medium from human VEGF-C baculovirus
25   infected insect cells, or pervanadate (a tyrosyl
phosphatase inhibitor) were used. As can be seen from
Fig. 33, in response to human baculoviral VEGF-C as well
as pervanadate treatment, VEGFR-2 was prominently
phosphorylated, whereas human and mouse recombinant VEGF-
30   C gave a weak and barely detectable enhancement of
autophosphorylation, respectively. Media from cell
cultures transfected with empty vector or VEGF-C cloned
in the antisense orientation did not induce
autophosphorylation of VEGFR-2. Therefore, mouse VEGF-C
35   binds to VEGFR-3 and activates this receptor at a much
lower concentration than needed for the activation of
VEGFR-2. Nevertheless, the invention comprehends methods

- 83 -

for using the materials of the invention to take
advantage of the interaction of VEGF-C with VEGFR-2, in
addition to the interaction between VEGF-C and VEGFR-3.

Example 27

5   **VEGF-C E104-S213 fragment expressed in Pichia yeast
stimulates autophosphorylation of Flt4 (VEGFR-3) and KDR
(VEGFR-2)**

         A truncated form of human VEGF-C cDNA was
constructed wherein (1) the sequence encoding residues of
10   a putative mature VEGF-C amino terminus H_2N-E(104)ETIK
(SEQ ID NO: 33, residues 104 et seq.) was fused in-frame
to the yeast PHO1 signal sequence (Invitrogen Pichia
Expression Kit, Catalog #K1710-01), and (2) a stop codon
was introduced after amino acid 213 (H_2N- ...RCMS; *i.e.,*
15   after codon 213 of SEQ ID NO: 32). The resultant
truncated cDNA construct was then inserted into the
Pichia pastoris expression vector pHIL-S1 (Invitrogen).
For the cloning, an internal *Bgl*II site in the VEGF-C
coding sequence was mutated without change of the encoded
20   polypeptide sequence.

         This VEGF-C expression vector was then
transfected into Pichia cells and positive clones were
identified by screening for the expression of VEGF-C
protein in the culture medium by Western blotting. One
25   positive clone was grown in a 50 ml culture, and induced
with methanol for various periods of time from 0 to 60
hours. About 10 μl of medium was analyzed by gel
electrophoresis, followed by Western blotting and
detection with anti-VEGF-C antiserum, as described above.
30   As can be seen in Figure 25, an approximately 24 kD
polypeptide (note the band spreading due to
glycosylation) accumulates in the culture medium of cells
transfected with the recombinant VEGF-C construct, but
not in the medium of mock-transfected cells or cells
35   transfected with the vector alone.

         The medium containing the recombinant VEGF-C
protein was concentrated by Centricon 30 kD cutoff

ultrafiltration and used to stimulate NIH 3T3 cells
expressing Flt4 (VEGFR-3) and porcine aortic endothelial
(PAE) cells expressing KDR (VEGFR-2). The stimulated
cells were lysed and immunoprecipitated using VEGFR-
5   specific antisera and the immunoprecipitates were
analyzed by Western blotting using anti-phosphotyrosine
antibodies, chemiluminescence, and fluorography. As a
positive control for maximal autophosphorylation of the
VEGFRs, vanadate (VO₄) treatment of the cells for 10
10   minutes was used. As can be seen from the results shown
in Fig. 26, medium from *Pichia* cultures secreting the
recombinant VEGF-C polypeptide induces
autophosphorylation of both Flt41 polypeptides of 195 kD
and 125 kD as well as the KDR polypeptide of about 200
15   kD. Vanadate, on the other hand, induces heavy tyrosyl
phosphorylation of the receptor bands in addition to
other bands probably coprecipitating with the receptors.
        These results demonstrate that a VEGF-
homologous domain of VEGF-C consisting of amino acid
20   residues 104E - 213S (SEQ ID NO: 33, residues 104-213)
can be recombinantly produced in yeast and is capable of
stimulating the autophosphorylation of Flt4 (VEGFR-3) and
KDR (VEGFR-2). Recombinant VEGF-C fragments such as the
fragment described herein, which are capable of
25   stimulating Flt4 or KDR autophosphorylation are intended
as aspects of the invention; methods of using these
fragments are also within the scope of the invention.

Example 28
Properties of the differentially
30   **processed forms of VEGF-C**
        The following oligonucleotides were used to
generate a set of VEGF-C variants:
        5'- TCTCTTCTGTGCTTGAGTTGAG -3' (SEQ ID NO: 42), used
to generate VEGF-C R102S (arginine mutated to serine at
35   position 102 (SEQ ID NO: 33));
        5'-TCTCTTCTGTCCCTGAGTTGAG -3' (SEQ ID NO: 43), used
to generate VEGF-C R102G (arginine mutated to glycine at

position 102 (SEQ ID NO: 33));

        5'-TGTGCTGCAGCAAATTTTATAGTCTCTTCTGTGGCGGCGGC
GGCGGCGGGCGCCTCGCGAGGACC -3' (SEQ ID NO: 44), used to
generate VEGF-C ΔN (deletion of N-terminal propeptide
corresponding to amino acids 32-102 (SEQ ID NO: 33));

        5'- CTGGCAGGGAACTGCTAATAATGGAATGAA - 3' (SEQ ID NO:
45), used to generate VEGF-C R226,227S (arginine codons
mutated to serines at positions 226 and 227 (SEQ ID NO:
33));

        5'-GGGCTCCGCGTCCGAGAGGTCGAGTCCGGACTCGTGATGGT
GATGGTGATGGGCGGCGGCGGCGGCGGGCGCCTCGCGAGGACC -3' (SEQ ID
NO: 46), used to generate VEGF-C NHis (this construct
encodes a polypeptide with a 6xHis tag fused to the N-
terminus of the secreted precursor (amino acid 32 of SEQ
ID NO: 33)).

        Some of the foregoing VEGF-C variant constructs
were further modified to obtain additional constructs.
For example, VEGF-C R102G in pALTER (Promega) and
oligonucleotide 5'-GTATTATAATGTCCTCCACCAAATTTTATAG -3'
(SEQ ID NO: 47) were used to generate VEGF-C 4G, which
encodes a polypeptide with four point mutations: R102G,
A110G, A111G, and A112G (alanines mutated to glycines at
positions 110-112 (SEQ ID NO: 33). These four mutations
are adjacent to predicted sites of cleavage of VEGF-C
expressed in PC-3 and recombinantly expressed in 293 EBNA
cells.

        Another construct was created using VEGF-C ΔN
and oligonucleotide 5'-GTTCGCTGCCTGACACTGTGGTAGTGTTGCTGGC
GGCCGCTAGTGATGGTGATGGTGATGAATAATGGAATGAACTTGTCTGTAAACATCC
AG -3' (SEQ ID NO: 48) to generate VEGF-C ΔNΔCHis. This
construct encodes a polypeptide with a deleted N-terminal
propeptide (amino acids 32-102); a deleted C-terminal
propeptide (amino acids 226-419 of SEQ ID NO: 33); and an
added 6xHis tag at the C-terminus.

        All constructs were further digested with
HindIII and NotI, subcloned into HindIII/NotI digested
pREP7 vector, and used to transfect 293 EBNA cells.

- 86 -

About 48 hours after transfection, the cells were either
metabolically labelled with Pro-mix™ as described above,
or starved in serum-free medium for 2 days. Media were
then collected and used in subsequent experiments. As

5   can be seen from Figs. 27A-B, wild type (wt) VEGF-C,
VEGF-C NHis and VEGF-C ΔNΔCHis were expressed to similar
levels in 293 EBNA cells. At the same time, expression
of the VEGF-C 4G polypeptide was considerably lower,
possibly due to the changed conformation and decreased

10   stability of the translated product. However, all the
above VEGF-C variants were secreted from the cells
(compare Figs. 27A and 27B). The conditioned media from
the transfected and starved cells were concentrated 5-
fold and used to assess their ability to stimulate

15   tyrosine phosphorylation of Flt4 (VEGFR-3) expressed in
NIH 3T3 cells and KDR (VEGFR-2) expressed in PAE cells.

        Figs. 28A-B show that wild type (wt) VEGF-C, as
well as all three mutant polypeptides, stimulate tyrosine
phosphorylation of VEGFR-3. The most prominent

20   stimulation is by the short mature VEGF-C ΔNΔCHis. This
mutant, as well as VEGF-C NHis, also stimulated tyrosine
phosphorylation of VEGFR-2. Thus, despite the fact that
a major component of secreted recombinant VEGF-C is a
dimer of 32/29 kD, the active part of VEGF-C responsible

25   for its binding to VEGFR-3 and VEGFR-2 is localized
between amino acids 102 and 226 (SEQ ID NO: 33) of the
VEGF-C precursor. Analysis and comparison of binding
properties and biological activities of these VEGF-C
proteins and variants, using assays such as those

30   described herein, will provide data concerning the
significance of the observed major 32/29 kD and 21-23 kD
VEGF-C processed forms. The data indicate that
constructs encoding amino acid residues 103-225 of the
VEGF-C precursor (SEQ ID NO: 33) generate a recombinant

35   ligand that is functional for both VEGFR-3 and VEGFR-2.
        The data from this and preceding examples
demonstrate that numerous fragments of the VEGF-C

polypeptide retain biological activity. A naturally
occurring VEGF-C polypeptide spanning amino acids 103-226
(or 103-227) of SEQ ID NO: 33, produced by a natural
processing cleavage defining the C-terminus, has been
5   shown to be active. Example 27 demonstrates that a
fragment with residues 104-213 of SEQ ID NO: 33 retains
biological activity.

In addition, data from Example 21 demonstrates
that a VEGF-C polypeptide having its amino terminus at
10   position 112 of SEQ ID NO: 33 retains activity.
Additional experiments have shown that a fragment lacking
residues 1-112 of SEQ ID NO: 33 retains biological
activity.

In a related experiment, a stop codon was
15   substituted for the lysine at position 214 of SEQ ID NO:
33 (SEQ ID NO: 32, nucleotides 991-993). The resulting
recombinant polypeptide still was capable of inducing
Flt4 autophosphorylation, indicating that a polypeptide
spanning amino acid residues 113-213 of SEQ ID NO: 33 is
20   biologically active.

Sequence comparisons of members of the VEGF
family of polypeptides provides an indication that still
smaller fragments of the polypeptide depicted in SEQ ID
NO: 33 will retain biological activity. In particular,
25   eight highly conserved cysteine residues of the VEGF
family of polypeptides define a region from residues 131 -
 211 of SEQ ID NO: 33 (see Figure 31) of evolutionary
signficance; therefore, a polypeptide spanning from about
residue 131 to about residue 211 is expected to retain
30   VEGF-C biological activity. In fact, a polypeptide which
retains the conserved motif RCXXCC (e.g., a polypeptide
comprising from about residue 161 to about residue 211 of
SEQ ID NO: 33 is postulated to retain VEGF-C biological
activity. To maintain native conformation of these
35   fragments, it may be preferred to retain about 1-2
additional amino acids at the carboxy-terminus and 1-2 or
more amino acids at the amino terminus.

Beyond the preceding considerations, evidence
exists that smaller fragments and/or fragment variants
which lack the conserved cysteines nonetheless will
retain VEGF-C biological activity. Consequently, the
5   materials and methods of the invention include all VEGF-C
fragments and variants that retain at least one
biological activity of VEGF-C, regardless of the presence
or absence of members of the conserved set of cysteine
residues.

10   **Example 29**
      **Expression of human VEGF-C under the human K14 keratin
      promoter in transgenic mice induces abundant growth of
                     lymphatic vessels in the skin**
                The Flt4 receptor tyrosine kinase is relatively
15   specifically expressed in the endothelia of lymphatic
vessels. Kaipainen et al., Proc. Natl. Acad. Sci. (USA),
92: 3566-3570 (1995). Furthermore, the VEGF-C growth
factor stimulates the Flt4 receptor, showing less
activity towards the KDR receptor of blood vessels
20   (Joukov et al., EMBO J., 15: 290-298 (1996); See Example
26).

                Experiments were conducted in transgenic mice
to analyze the specific effects of VEGF-C overexpression
in tissues. The human K14 keratin promoter is active in
25   the basal cells of stratified squamous epithelia (Vassar
et al., Proc. Natl. Acad. Sci. (USA), 86:1563-1567
(1989)) and was used as the expression control element in
the recombinant VEGF-C transgene. The vector containing
the K14 keratin promoter is described in Vassar et al.,
30   Genes Dev., 5:714-727 (1991) and Nelson et al., J. Cell
Biol. 97:244-251 (1983).

                The recombinant VEGF-C transgene was
constructed using the human full length VEGF-C cDNA
(GenBank Acc. No. X94216). This sequence was excised
35   from a pCI-neo vector (Promega) with XhoI/NotI, and the
resulting 2027 base pair fragment containing the open
reading frame and stop codon (nucleotides 352-1611 of SEQ

ID NO: 32) was isolated. The isolated fragment was then
subjected to an end-filling reaction using the Klenow
fragment of DNA polymerase. The blunted fragment was
then ligated to a similarly opened *Bam*HI restriction site
5   in the K14 vector. The resulting construct contained the
*Eco*RI site derived from the polylinker of the pCI-neo
vector. This *Eco*RI site was removed using standard
techniques (a Klenow-mediated fill-in reaction following
partial digestion of the recombinant intermediate with
10   *Eco*RI) to facilitate the subsequent excision of the DNA
fragment to be injected into fertilized mouse oocytes.
The resulting clone, designated K14-VEGF-C, is
illustrated in Fig. 20.

          The *Eco*RI-*Hind*III fragment from clone K14 VEGF-
15   C containing the K14 promoter, VEGF-C cDNA, and K14
polyadenylation signal was isolated and injected into
fertilized oocytes of the FVB-NIH mouse strain. The
injected zygotes were transplanted to oviducts of
pseudopregnant C57BL/6 x DBA/2J hybrid mice. The
20   resulting founder mice were analyzed for the presence of
the transgene by polymerase chain reaction of tail DNA
using the primers: 5'-CATGTACGAACCGCCAG-3' (SEQ ID NO:
49) and 5'-AATGACCAGAGAGAGGCGAG-3' (SEQ ID NO: 50). In
addition, the tail DNAs were subjected to *Eco*RV digestion
25   and subsequent Southern analysis using the *Eco*RI-*Hind*III
fragment injected into the mice. Out of 8 pups analyzed
at 3 weeks of age, 2 were positive, having approximately
40-50 copies and 4-6 copies of the transgene in their
respective genomes.

30   The mouse with the high copy number transgene
was small, developed more slowly than its litter mates
and had difficulty eating (*i.e.*, suckling). Further
examination showed a swollen, red snout and poor fur.
Although fed with a special liquid diet, it suffered from
35   edema of the upper respiratory and digestive tracts after
feeding and had breathing difficulties. This mouse died
eight weeks after birth and was immediately processed for

- 90 -

histology, immunohistochemistry, and *in situ*
hybridization.

Histological examination showed that in
comparison to the skin of littermates, the dorsal dermis
of K14-VEGF-C transgenic mice was atrophic and connective
tissue was replaced by large lacunae devoid of red cells,
but lined with a thin endothelial layer (white arrows in
Figs. 29A-D). These distended vessel-like structures
resembled those seen in human lymphangiomas. The number
of skin adnexal organs and hair follicles were reduced.
In the snout region, an increased number of vessels was
also seen. Therefore, VEGF-C overexpression in the basal
epidermis is capable of promoting the growth of extensive
vessel structure in the underlying skin, including large
vessel lacunae. The endothelial cells surrounding these
lacunae contained abundant Flt4 mRNA in *in situ*
hybridization (see Examples 23 and 30 for methodology).
The vessel morphology indicates that VEGF-C stimulates
the growth of vessels having features of lymphatic
vessels. The other K14-VEGF-C transgenic mouse had a
similar skin histopathology.

The foregoing *in vivo* data indicates utilities
for both (i) VEGF-C polypeptides and polypeptide variants
having VEGF-C biological activity, and (ii) anti-VEGF-C
antibodies and VEGF-C antagonists that inhibit VEGF-C
activity (e.*g.*, by binding VEGF-C or interfering with
VEGF-C/receptor interactions. For example, the data
indicates a therapeutic utility for VEGF-C polypeptides
in patients wherein growth of lymphatic tissue may be
desirable (e.g., in patients following breast cancer or
other surgery where lymphatic tissue has been removed and
where lymphatic drainage has therefore been compromised,
resulting in swelling; or in patients suffering from
elephantiasis). The data indicates a therapeutic utility
for anti-VEGF-C antibody substances and VEGF-C
antagonists for conditions wherein growth-inhibition of
lymphatic tissue may be desirable (e.g., treatment of

- 91 -

lymphangiomas). Accordingly, methods of administering
VEGF-C and VEGF-C variants and antagonists are
contemplated as methods and materials of the invention.

Example 30
Expression of VEGF-C and Flt4
in the Developing Mouse

Embryos from a 16-day post-coitus pregnant
mouse were prepared and fixed in 4% paraformaldehyde
(PFA), embedded in paraffin, and sectioned at 6 μm. The
sections were placed on silanated microscope slides and
treated with xylene, rehydrated, fixed for 20 minutes in
4% PFA, treated with proteinase K (7mg/ml; Merck,
Darmstadt, Germany) for 5 minutes at room temperature,
again fixed in 4% PFA and treated with acetic anhydride,
dehydrated in solutions with increasing ethanol
concentrations, dried and used for in situ hybridization.

In situ hybridization of sections was performed
as described (Västrik et al., J. Cell Biol.,
128:1197-1208 (1995)). A mouse VEGF-C antisense RNA
probe was generated from linearized pBluescript II SK+
plasmid (Stratagene Inc.), containing a fragment
corresponding to nucleotides 499-979 of mouse VEGF-C
cDNA, where the noncoding region and the BR3P repeat were
removed by Exonuclease III treatment. The fragment had
been cloned into the EcoRI and HindIII sites of
pBluescript II SK+. Radiolabeled RNA was synthesized
using T7-RNA Polymerase and [^{35}S]-UTP (Amersham, Little
Chalfont, UK). About two million cpm of the VEGF-C probe
was applied per slide. After an overnight hybridization,
the slides were washed first in 2x SSC and 20-30 mM DDT
for 1 hour at 50°C. Treatment continued with a high
stringency wash, 4x SSC and 20 mM DTT and 50% deionized
formamide for 30 minutes at 65°C followed by RNase A
treatment (20 μg/ml) for 30 minutes at 37°C. The high
stringency wash was repeated for 45 minutes. Finally,
the slides were dehydrated and dried for 30 minutes at

- 92 -

room temperature. The slides were dipped into
photography emulsion and exposed for 4 weeks. Slides
were developed using Kodak D-16 developer, counterstained
with hematoxylin and mounted with Permount
5   (FisherChemical).

For *in situ* hybridizations of Flt4 sequences, a
mouse Flt4 cDNA fragment covering bp 1-192 of the
published sequence (Finnerty *et al.*, *Oncogene*, 8:2293-
2298 (1993)) was used, and the above-described protocol
10 was followed, with the following exceptions.
Approximately one million cpm of the Flt4 probe were
applied to each slide. The stringent washes following
hybridization were performed in 1x SSC and 30 mM DTT for
105 minutes.

15   The figure shows photomicrographs of the
hybridized sections in dark field microscopy (Figs. 36A-
C) and light field microscopy (Fig. 36D). Magnifications
used for photography were 4x for Figs. 36A-B and 10x for
Figs. 36C-D. The transverse sections shown are from the
20 cephalic region and the area shown for VEGF-C and FLT4
are about 14 sections apart, Flt4 being more cranially
located in the embryo. In Fig. 36A (Flt4 probe), the
developing nasopharyngeal cavity is in the midline in the
upper, posterior part; in the anterior part of Fig. 36A
25 is the snout with emerging fibrissal follicles and, in
the midline, the forming nasal cavity. On both sides,
the retinal pigment gives a false positive signal in dark
field microscopy. The most prominently Flt4-hybridizing
structures appear to correspond to the developing
30 lymphatic and venous endothelium. Note that a
plexus-like endothelial vascular structure surrounds the
developing nasopharyngeal mucous membrane. In Fig. 36B,
the most prominent signal is obtained from the posterior
part of the developing nasal conchae, which in higher
35 magnification (Figs. 36C-D) show the epithelium
surrounding loose connective tissue/forming cartilage.
This structure gives a strong *in situ* hybridization

signal for VEGF-C. Also in Fig. 36B, more weakly
hybridizing areas can be seen around the snout, although
this signal is much more homogeneous in appearance.
Thus, the expression of VEGF-C is strikingly high in the
5   developing nasal conchae.

        The conchae are surrounded with a rich vascular
plexus, important in nasal physiology as a source for the
mucus produced by the epithelial cells and for warming
inhaled air. It is suggested that VEGF-C is important in
10   the formation of the concheal venous plexus at the mucous
membranes, and that it may also regulate the permeability
of the vessels needed for the secretion of nasal mucus.
Possibly, VEGF-C and its derivatives, and antagonists,
could be used in the regulation of the turgor of the
15   conchal tissue and mucous membranes and therefore the
diameter of the upper respiratory tract, as well as the
quantity and quality of mucus produced. These factors
are of great clinical significance in inflammatory
(including allergic) and infectious diseases of the upper
20   respiratory tract. Accordingly, the invention
contemplates the use of the materials of the invention,
including VEGF-C, Flt4, and their derivatives, in methods
of diagnosing and treating inflammatory and infectious
conditions affecting the upper respiratory tract,
25   including nasal structures.

                        Example 31
        Characterization of the exon-intron organization
                    of the human VEGF-C gene
        Two genomic DNA clones covering exons 1, 2, and
30   3 of the human VEGF-C gene were isolated from a human
genomic DNA library using VEGF-C cDNA fragments as
probes. In particular, a human genomic library in
bacteriophage EMBL-3 lambda (Clontech) was screened using
a PCR-generated fragment corresponding to nucleotides
35   629-746 of the human VEGF-C cDNA (SEQ ID NO: 32). One
positive clone, designated "lambda 3," was identified,
and the insert was subcloned as a 14 kb *Xho*I fragment

- 94 -

into the pBSK II vector (Stratagene). The genomic
library also was screened with a labeled 130 bp *NotI-SacI*
fragment from the 5'-noncoding region of the VEGF-C cDNA
(the *NotI* site is in the polylinker of the cloning

5   vector; the *SacI* site corresponds to nucleotides 92-97 of
SEQ ID NO: 32). Two positive clones, designated "lambda
5" and "lambda 8," were obtained. Restriction mapping
analysis showed that clone lambda 3 contains exons 2 and
3, while clone lambda 5 contains exon 1 and the putative

10   promoter region.

        Three genomic fragments containing exons 4, 5,
6 and 7 were subcloned from a genomic VEGF-C P1 plasmid
clone. In particular, purified DNA from a genomic P1
plasmid clone 7660 (Paavonen et al., *Circulation, 93*:

15   1079-1082 (1996)) was used. *EcoRI* fragments of the P1
insert DNA were ligated into pBSK II vector. Subclones
of clone 7660 which contained human VEGF-C cDNA
homologous sequences were identified by colony
hybridization, using the full-length VEGF-C cDNA as a

20   probe. Three different genomic fragments were identified
and isolated, which contained the remaining exons 4-7.

        To determine the genomic organization, the
clones were mapped using restriction endonuclease
cleavage. Also, the coding regions and exon-intron

25   junctions were partially sequenced. The result of this
analysis is depicted in Figures 11 and 17. The sequences
of all intron-exon boundaries (Fig. 17, SEQ ID NOs: 57-
68) conformed to the consensus splicing signals (Mount,
Nucl. Acids Res., 10: 459-472 (1982)). The length of the

30   intron between exon 5 and 6 was determined directly by
nucleotide sequencing and found to be 301 bp. The length
of the intron between exons 2 and 3 was determined by
restriction mapping and Southern hybridization and was
found to be about 1.6 kb. Each of the other introns was

35   over 10 kb in length.

        A similar analysis was performed for the murine
genomic VEGF-C gene. The sequences of murine VEGF-C

intron-exon boundaries are depicted in Figure 17 and SEQ
ID NOs: 69-80.

     The restriction mapping and sequencing data
indicated that the signal sequence and the first residues
5   of the N-terminal propeptide are encoded by exon 1. The
second exon encodes the carboxy-terminal portion of the
N-terminal propeptide and the amino terminus of the VEGF
homology domain. The most conserved sequences of the
VEGF homology domain are distributed in exons 3
10   (containing 6 conserved cysteine residues) and 4
(containing 2 cys residues). The remaining exons encode
cysteine-rich motifs of the type C-6X-C-10X-CRC (exons 5
and 7) and a fivefold repeated motif of type
C-6X-B-3X-C-C-C, which is typical of a silk protein.
15   To further characterize the VEGF-C gene
promoter, the lambda 5 clone was further analyzed.
Restriction mapping of this clone using a combination of
single- and double-digestions and Southern hybridizations
indicated that it includes: (1) an approximately 5 kb
20   region upstream of the putative initiator ATG codon, (2)
exon 1, and (3) part of intron I of the VEGF-C gene.

    A 3.7 kb *Xba* I fragment of clone lambda 5,
containing exon 1 and 5' and 3' flanking sequences, was
subcloned and further analyzed. As reported previously,
25   a major VEGF-C mRNA band migrates at a position of about
2.4 kb. Calculating from the VEGF-C coding sequence of
1257 bp and a 391 bp 3' noncoding sequence plus a polyA
sequence of about 50-200 bp, the mRNA start site should
be located about 550-700 bp upstream of the translation
30   initiation codon.

    To further characterize the promoter of the
human VEGF-C gene, a genomic clone encompassing about 1.4
kb upstream of the translation initiation site was
isolated, and the 5' noncoding cDNA sequence and putative
35   promoter region were sequenced. The sequence obtained is
set forth in SEQ ID NO: 54. Similar to what has been
observed with the VEGF gene, the VEGF-C promoter is rich

- 96 -

in G and C residues and lacks consensus TATA and CCAAT
sequences. Instead, it has numerous putative binding
sites for Sp1, a ubiquitous nuclear protein that can
initiate transcription of TATA-less genes. See Pugh and
5   Tjian, *Genes and Dev.*, 5:105-119 (1991). In addition,
sequences upstream of the VEGF-C translation start site
were found to contain frequent consensus binding sites
for the AP-2 factor. This suggests that the
cAMP-dependent protein kinase and protein kinase C, as
10   activators of AP-2 transcription factor [Curran and
Franza, *Cell*, 55:395-397 (1988)], mediate VEGF-C
transcriptional regulation.

      The VEGF-C gene is abundantly expressed in
adult human tissues, such as heart, placenta, ovary and
15   small intestine, and is induced by a variety of factors.
Indeed, several potential binding sites for regulators of
tissue-specific gene expression, like NFkB and GATA, are
located in the distal part of the VEGF-C promoter. For
example, NFkB is known to regulate the expression of
20   tissue factor in endothelial cells. Also, transcription
factors of the GATA family are thought to regulate
cell-type specific gene expression.

      Unlike VEGF, the VEGF-C gene does not contain a
binding site for the hypoxia-inducible factor, HIF-1
25   (Levy et al., *J. Biol. Chem.*, 270: 13333-13340 (1995)).
This finding suggests that if the VEGF-C mRNA is
regulated by hypoxia, the mechanism would be based mainly
on the regulation of mRNA stability. In this regard,
numerous studies have shown that the major control point
30   for the hypoxic induction of the VEGF gene is the
regulation of the steady-state level of mRNA. See Levy
et al., *J. Biol. Chem.*, 271: 2746-2753 (1996). The
relative rate of VEGF mRNA stability and decay is
considered to be determined by the presence of specific
35   sequence motifs in its 3' untranslated region (UTR),
which have been demonstrated to regulate mRNA stability
(Chen and Shyu, *Mol. Cell Biol.*, 14: 8471-8482 (1994)).
The 3'-UTR of the VEGF-C gene also contains a putative

motif of this type (TTATTT), at positions 1873-1878 of
SEQ ID NO: 32.

Example 32
Identification of a VEGF-C splice variant

5   As reported in Example 16, a major 2.4 kb VEGF-
C mRNA and smaller amounts of a 2.0 kb mRNA are
observable. To clarify the origin of these RNAs, several
additional VEGF-C cDNAs were isolated and characterized.
A human fibrosarcoma cDNA library from HT1080 cells in
10   the lambda gt11 vector (Clontech, product #HL1048b) was
screened using a 153 bp human VEGF-C cDNA fragment as a
probe as described in Example 10. See also Joukov et
al., EMBO J., 15:290-298 (1996). Nine positive clones
were picked and analyzed by PCR amplification using
15   oligonucleotides 5'-CACGGCTTATGCAAGCAAAG-3' (SEQ ID NO:
55) and 5'-AACACAGTTTTCCATAATAG-3' (SEQ ID NO: 56) These
oligonucleotides were selected to amplify the portion of
the VEGF-C cDNA corresponding to nucleotides 495-1661 of
SEQ ID NO: 32. PCR was performed using an annealing
20   temperature of 55°C and 25 cycles.

            The resultant PCR products were electrophoresed
on agarose gels. Five clones out of the nine analyzed
generated PCR fragments of the expected length of 1147
base pairs, whereas one was slightly shorter. The
25   shorter fragment and one of the fragments of expected
length were cloned into the pCRTMII vector (Invitrogen)
and analyzed by sequencing. The sequence revealed that
the shorter PCR fragment had a deletion of 153 base
pairs, corresponding to nucleotides 904 to 1055 of SEQ ID
30   NO: 32. These deleted bases correspond to exon 4 of the
human and mouse VEGF-C genes, schematically depicted in
Fig. 17. Deletion of exon 4 results in a frameshift,
which in turn results in a C-terminal truncation of the
full-length VEGF-C precursor, with fifteen amino acid
35   residues translated from exon 5 in a different frame than
the frame used to express the full-length protein. Thus,

the C-terminal amino acid sequence of the resulting
truncated polypeptide would be --Leu (181)-Ser-Lys-Thr-
Val-Ser-Gly-Ser-Glu-Gln-Asp-Leu-Pro-His-Glu-Leu-His-Val-
Glu(199) (SEQ ID NO: 81). The VEGF-C variant encoded by
5   this splice variant would not contain the C-terminal
cleavage site of the VEGF-C precursor. Thus, a putative
alternatively spliced RNA form lacking conserved exon 4
was identified in HT-1080 fibrosarcoma cells and this
form is predicted to encode a protein of 199 amino acid
10   residues, which could be an antagonist of VEGF-C.

                        Example 33
VEGF-C is similarly processed in different cell cultures
in vitro
15   To study whether VEGF-C is similarly processed
in different cell types, 293 EBNA cells, COS-1 cells and
HT-1080 cells were transfected with wild type human VEGF-
C cDNA and labelled with Pro-Mix™ as described in Exam-
ple 22. The conditioned media from the cultures were col-
20   lected and subjected to immunoprecipitation using an-
tiserum 882 (described in Example 21, recognizing a pep-
tide corresponding to amino acids 104-120 of SEQ ID NO:
33). The immunoprecipitated polypeptides were separated
via SDS-PAGE, and detected via autoradiography. The major
25   form of secreted recombinant VEGF-C observed from all
cell lines tested is a 29/32 kD doublet. These two poly-
peptides are bound to each other by disulfide bonds, as
described in Example 22. A less prominent band of appro-
ximately 21 kD also was detected in the culture media.
30   Additionally, a non-processed VEGF-C precursor of 63 kDa
was observed. This form was more prominent in the COS-1
cells, suggesting that proteolytic processing of VEGF-C
in COS cells is less efficient than in 293 EBNA cells.
Endogenous VEGF-C (in non-transfected cells) was not de-
35   tectable under these experimental conditions in the
HT-1080 cells, but was readily detected in the condi-
tioned medium of the PC-3 cells. Analysis of the subunit

polypeptide sizes and ratios in PC-3 cells and 293 EBNA
cells revealed strikingly similar results: the most pro-
minent form was a doublet of 29/32 kDa and a less pro-
minent form the 21 kD polypeptide. The 21 kD form produ-
5   ced by 293 EBNA cells was not recognized by the 882 anti-
body in the Western blot, although it is recognized when
the same antibody is used for immunoprecipitation (see
data in previous examples). As reported in Example 21,
cleavage of the 32 kD form in 293 EBNA cells occurs bet-
10   ween amino acid residues 111 and 112 (SEQ ID NO: 33),
downstream of the cleavage site in PC-3 cells (between
residues 102 and 103). Therefore, the 21 kD form produced
in 293 EBNA cells does not contain the complete N-ter-
minal peptide used to generate antiserum 882.
15   In a related experiment, PC-3 cells were
cultured in serum-free medium for varying periods of time
(1 - 8 days) prior to isolation of the conditioned me-
dium. The conditioned medium was concentrated using a
Centricon device (Amicon, Beverly, USA) and subjected to
20   Western blotting analysis using antiserum 882. After one
day of culturing, a prominent 32 kD band was detected.
Increasing amounts of a 21-23 kD form were detected in
the conditioned media from 4 day and 8 day cultures. The
diffuse nature of this polypeptide band, which is simply
25   called the 23 kD polypeptide in example 5 and several
subsequent examples is most likely due to a heterogenous
and variable amount of glycosylation. These results indi-
cate that, initially, the cells secrete a 32 kD polypep-
tide, which is further processed or cleaved in the medium
30   to yield the 21-23 kD form. The microheterogeneity of
this polypeptide band would then arise from the variable
glycosylation degree and, from microheterogeneity of the
processing cleavage sites, such as obtained for the amino
terminus in PC-3 and 293 EBNA cell cultures. The carboxyl
35   terminal cleavage site could also vary, examples of pos-
sible cleavage sites would be between residues 225-226,
226-227 and 227-228 as well as between residues 216-217.

- 100 -

Taken together, these data suggest the possibility that secreted cellular protease(s) are responsible for the generation of the 21-23 kD form of VEGF-C from the 32 kD polypeptide. Such proteases could be used in vitro to

5   cleave VEGF-C precursor proteins in solution during the production of VEGF-C, or used in cell culture and in vivo to release biologically active VEGF-C.

Example 34

10   **Differential binding of VEGF-C forms by the extracellular domains of VEGFR-3 and VEGFR-2**

In two parallel experiments, 293 EBNA cells were transfected with a construct encoding recombinant wild type VEGF-C or a construct encoding VEGF-C DNDCHis

15   (Example 28) and about 48 hours after transfection, metabolically labelled with Pro-Mix™ as described in previous examples. The media were collected from mock-transfected and transfected cells and used for receptor binding analyses.

20   Receptor binding was carried out in binding buffer (PBS, 0.5% BSA, 0.02% Tween 20, 1 microgram/ml heparin) containing approximately 0.2 microgram of either (a) a fusion protein comprising a VEGFR-3 extracellular domain fused to an immunoglobulin sequence (VEGFR-3-Ig)

25   or (b) a fusion protein comprising VEGFR-2 extracellular domain fused to an alkaline phosphatase sequence (VEGF-R-2-AP; Cao et al., J. Biol. Chem. 271:3154-62 (1996)). As a control, similar aliquots of the 293 EBNA conditioned media were mixed with 2 µl of anti-VEGF-C an-

30   tiserum (VEGF-C IP).

After incubation for 2 hours at room temperature, anti-VEGF-C antibodies and VEGFR-3-Ig protein were adsorbed to protein A-sepharose (PAS) and VEGFR-2-AP was immunoprecipitated using anti-AP monoclonal antibodies

35   (Medix Biotech, Genzyme Diagnostics, San Carlos, CA, USA) and protein G-sepharose. Complexes containing VEGF-C bound to VEGFR-3-Ig or VEGFR-2-AP were washed three times

- 101 -

in binding buffer, twice in 20 mM Tris-HCl (pH 7.4) and
VEGF-C immunoprecipitates were washed three times in RIPA
buffer and twice in 20 mM tris-HCl (pH 7.4) and analyzed
via SDS-PAGE under reducing and nonreducing conditions.
5   As a control, the same media were precipitated with anti-
AP and protein G-sepharose (PGS) or with PAS to control
for possible nonspecific adsorption.

         These experiments revealed that VEGFR-3 bound
to both the 32/29 kD and 21-23 kD forms of recombinant
10 VEGF-C, whereas VEGFR-2 bound preferentially to the 21-23
kD component from the conditioned media. In addition,
small amounts of 63 kD and 52 kD VEGF-C forms were obser-
ved binding with VEGFR-3. Further analysis under nonredu-
cing conditions indicates that a great proportion of the
15 21-23 kD VEGF-C bound to either receptor does not contain
interchain disulfide bonds. These findings reinforce the
results that VEGF-C binds VEGFR-2. This data suggests a
utility for recombinant forms of VEGF-C which are active
towards VEGFR-3 only or which are active towards both
20 VEGFR-3 and VEGFR-2. On the other hand, these results,
together with the results in Example 28, do not eliminate
the possibility that the 32/29 kD dimer binds VEGFR-3 but
does not activate it. The failure of the 32/29 kD dimer
to activate VEGFR-3 could explain the finding that condi-
25 tioned medium from the N-His VEGF-C transfected cells
induced a less prominent tyrosine phosphorylation of
VEGFR-3 than medium from VEGF-C DNDCHis transfected
cells, even though expression of the former polypeptide
was much higher (see Figs. 27 and 28). Stable VEGF-C po-
30 lypeptide variants that bind to a VEGF-C receptor but
fail to activate the receptor are useful as VEGF-C anta-
gonists.

Deposit of Biological Materials: Plasmid FLT4-
L has been deposited with the American Type Culture
Collection (ATCC), 12301 Parklawn Dr., Rockville MD 20952
5   (USA), pursuant to the provisions of the Budapest Treaty,
and has been assigned a deposit date of 24 July 1995 and
ATCC accession number 97231.

While the present invention has been described
in terms of specific embodiments, it is understood that
10   variations and modifications will occur to those in the
art.   Accordingly, only such limitations as appear in the
appended claims should be placed on the invention.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

    (i) APPLICANT: Helsinki University Licensing Ltd Oy

   (ii) TITLE OF INVENTION: Receptor Ligand VEGF-C

 (iii) NUMBER OF SEQUENCES: 81

   (iv) CORRESPONDENCE ADDRESS:
      (A) ADDRESSEE: Oy Jalo Ant-Wuorinen Ab
      (B) STREET: Iso Roobertinkatu 4-6 A
      (C) CITY: Helsinki
      (E) COUNTRY: Finland
      (F) ZIP: 00120

    (v) COMPUTER READABLE FORM:
      (A) MEDIUM TYPE: Floppy disk
      (B) COMPUTER: IBM PC compatible
      (C) OPERATING SYSTEM: PC-DOS/MS-DOS
      (D) SOFTWARE: PatentIn Release #1.0, Version #1.30

   (vi) CURRENT APPLICATION DATA:
      (A) APPLICATION NUMBER:
      (B) FILING DATE:
      (C) CLASSIFICATION:

  (vii) PRIOR APPLICATION DATA:
      (A) APPLICATION NUMBER: 08/671,573
      (B) FILING DATE: 28-JUN-1996

  (vii) PRIOR APPLICATION DATA:
      (A) APPLICATION NUMBER: 08/601,132
      (B) FILING DATE: 14-FEB-1996

  (vii) PRIOR APPLICATION DATA:
      (A) APPLICATION NUMBER: 08/585,895
      (B) FILING DATE: 12-JAN-1996

  (vii) PRIOR APPLICATION DATA:
      (A) APPLICATION NUMBER: 08/510,133
      (B) FILING DATE: 01-AUG-1995

 (viii) ATTORNEY/AGENT INFORMATION:
      (A) NAME: Karvinen, Leena
      (C) REFERENCE/DOCKET NUMBER: 28999

   (ix) TELECOMMUNICATION INFORMATION:
      (A) TELEPHONE: +358 0 648 606
      (B) TELEFAX: +358 0 640 575
      (C) TELEX: 123505 JALO FI

(2) INFORMATION FOR SEQ ID NO:1:

    (i) SEQUENCE CHARACTERISTICS:
      (A) LENGTH: 20 base pairs
      (B) TYPE: nucleic acid
      (C) STRANDEDNESS: single
      (D) TOPOLOGY: linear

   (ii) MOLECULE TYPE: cDNA

   (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

TGTCCTCGCT GTCCTTGTCT   20

(2) INFORMATION FOR SEQ ID NO:2:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 70 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: cDNA

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

ACATGCATGC CACCATGCAG CGGGGCGCCG CGCTGTGCCT GCGACTGTGG CTCTGCCTGG   60

GACTCCTGGA   70

(2) INFORMATION FOR SEQ ID NO:3:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 24 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: cDNA

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

ACATGCATGC CCCGCCGGTC ATCC   24

(2) INFORMATION FOR SEQ ID NO:4:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 22 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: cDNA

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

CGGAATTCCC CATGACCCCA AC   22

(2) INFORMATION FOR SEQ ID NO:5:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 33 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: cDNA

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

CCATCGATGG ATCCTACCTG AAGCCGCTTT CTT   33

(2) INFORMATION FOR SEQ ID NO:6:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 17 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: cDNA

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

ATTTAGGTGA CACTATA   17

(2) INFORMATION FOR SEQ ID NO:7:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 34 base pairs
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: single
            (D) TOPOLOGY: linear

     (ii) MOLECULE TYPE: cDNA

     (xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

CCATCGATGG ATCCCGATGC TGCTTAGTAG CTGT   34

(2) INFORMATION FOR SEQ ID NO:8:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 40 amino acids
            (B) TYPE: amino acid
            (C) STRANDEDNESS: single
            (D) TOPOLOGY: linear

     (ii) MOLECULE TYPE: peptide

     (xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

     Pro Met Thr Pro Thr Thr Tyr Lys Gly Ser Val Asp Asn Gln Thr Asp
     1   5   10   15

     Ser Gly Met Val Leu Ala Ser Glu Glu Phe Glu Gln Ile Glu Ser Arg
                 20   25   30

     His Arg Gln Glu Ser Gly Phe Arg
             35   40

(2) INFORMATION FOR SEQ ID NO:9:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 21 base pairs
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: single
            (D) TOPOLOGY: linear

     (ii) MOLECULE TYPE: cDNA

     (xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

CTGGAGTCGA CTTGGCGGAC T   21

(2) INFORMATION FOR SEQ ID NO:10:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 60 base pairs
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: single
            (D) TOPOLOGY: linear

     (ii) MOLECULE TYPE: cDNA

     (xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

CGCGGATCCC TAGTGATGGT GATGGTGATG TCTACCTTCG ATCATGCTGC CCTTATCCTC   60

(2) INFORMATION FOR SEQ ID NO:11:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 34 base pairs
            (B) TYPE: nucleic acid

```
            (C) STRANDEDNESS: single
            (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: cDNA

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

CCCAAGCTTG GATCCAAGTG GCTACTCCAT GACC                              34

(2) INFORMATION FOR SEQ ID NO:12:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 20 base pairs
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: single
            (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: cDNA

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

GTTGCCTGTG ATGTGCACCA                                             20

(2) INFORMATION FOR SEQ ID NO:13:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 18 amino acids
            (B) TYPE: amino acid
            (C) STRANDEDNESS: single
            (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

    Xaa Glu Glu Thr Ile Lys Phe Ala Ala Ala His Tyr Asn Thr Glu Ile
    1               5                   10                  15

    Leu Lys


(2) INFORMATION FOR SEQ ID NO:14:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 17 base pairs
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: single
            (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: cDNA

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

GCAGARGARA CNATHAA                                                17

(2) INFORMATION FOR SEQ ID NO:15:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 5 amino acids
            (B) TYPE: amino acid
            (C) STRANDEDNESS: single
            (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

    Glu Glu Thr Ile Lys
    1               5
```

(2) INFORMATION FOR SEQ ID NO:16:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 18 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: cDNA

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

GCAYTTNARD ATYTCNGT   18

(2) INFORMATION FOR SEQ ID NO:17:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 5 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

Thr Glu Ile Leu Lys
1   5

(2) INFORMATION FOR SEQ ID NO:18:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 22 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: cDNA

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

ATTCGCTGCA GCACACTACA AC   22

(2) INFORMATION FOR SEQ ID NO:19:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 19 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: cDNA

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

TCNGTGTTGT AGTGTGCTG   19

(2) INFORMATION FOR SEQ ID NO:20:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 7 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

Ala Ala His Tyr Asn Thr Glu

```
        1                5
```

(2) INFORMATION FOR SEQ ID NO:21:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 20 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: cDNA

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

TAATACGACT CACTATAGGG   20

(2) INFORMATION FOR SEQ ID NO:22:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 24 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: cDNA

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

GTTGTAGTGT GCTGCAGCGA ATTT   24

(2) INFORMATION FOR SEQ ID NO:23:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 8 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

    Lys Phe Ala Ala Ala His Tyr Asn
    1   5

(2) INFORMATION FOR SEQ ID NO:24:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 21 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: cDNA

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

TCACTATAGG GAGACCCAAG C   21

(2) INFORMATION FOR SEQ ID NO:25:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 219 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: cDNA

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

TCACTATAGG GAGACCCAAG CTTGGTACCG AGCTCGGATC CACTAGTAAC GGCCGCCAGT   60

GTGGTGGAAT TCGACGAACT CATGACTGTA CTCTACCCAG AATATTGGAA AATGTACAAG   120

TGTCAGCTAA GGCAAGGAGG CTGGCAACAT AACAGAGAAC AGGCCAACCT CAACTCAAGG   180

ACAGAAGAGA CTATAAAATT CGCTGCAGCA CACTACAAC   219

(2) INFORMATION FOR SEQ ID NO:26:

      (i) SEQUENCE CHARACTERISTICS:
          (A) LENGTH: 18 base pairs
          (B) TYPE: nucleic acid
          (C) STRANDEDNESS: single
          (D) TOPOLOGY: linear

     (ii) MOLECULE TYPE: cDNA

     (xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

ACAGAGAACA GGCCAACC   18

(2) INFORMATION FOR SEQ ID NO:27:

      (i) SEQUENCE CHARACTERISTICS:
          (A) LENGTH: 19 base pairs
          (B) TYPE: nucleic acid
          (C) STRANDEDNESS: single
          (D) TOPOLOGY: linear

     (ii) MOLECULE TYPE: cDNA

     (xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

TCTAGCATTT AGGTGACAC   19

(2) INFORMATION FOR SEQ ID NO:28:

      (i) SEQUENCE CHARACTERISTICS:
          (A) LENGTH: 25 base pairs
          (B) TYPE: nucleic acid
          (C) STRANDEDNESS: single
          (D) TOPOLOGY: linear

     (ii) MOLECULE TYPE: cDNA

     (xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

AAGAGACTAT AAAATTCGCT GCAGC   25

(2) INFORMATION FOR SEQ ID NO:29:

      (i) SEQUENCE CHARACTERISTICS:
          (A) LENGTH: 20 base pairs
          (B) TYPE: nucleic acid
          (C) STRANDEDNESS: single
          (D) TOPOLOGY: linear

     (ii) MOLECULE TYPE: cDNA

     (xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

CCCTCTAGAT GCATGCTCGA   20

(2) INFORMATION FOR SEQ ID NO:30:

      (i) SEQUENCE CHARACTERISTICS:
          (A) LENGTH: 24 base pairs
          (B) TYPE: nucleic acid
          (C) STRANDEDNESS: single

              (D) TOPOLOGY: linear

        (ii) MOLECULE TYPE: cDNA

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

GTTGTAGTGT GCTGCAGCGA ATTT   24

(2) INFORMATION FOR SEQ ID NO:31:

        (i) SEQUENCE CHARACTERISTICS:
              (A) LENGTH: 21 base pairs
              (B) TYPE: nucleic acid
              (C) STRANDEDNESS: single
              (D) TOPOLOGY: linear

        (ii) MOLECULE TYPE: cDNA

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

TCACTATAGG GAGACCCAAG C   21

(2) INFORMATION FOR SEQ ID NO:32:

        (i) SEQUENCE CHARACTERISTICS:
              (A) LENGTH: 1997 base pairs
              (B) TYPE: nucleic acid
              (C) STRANDEDNESS: single
              (D) TOPOLOGY: linear

        (ii) MOLECULE TYPE: cDNA

        (ix) FEATURE:
              (A) NAME/KEY: CDS
              (B) LOCATION: 352..1608

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

CCCGCCCCGC CTCTCCAAAA AGCTACACCG ACGCGGACCG CGGCGGCGTC CTCCCTCGCC   60

CTCGCTTCAC CTCGCGGGCT CCGAATGCGG GGAGCTCGGA TGTCCGGTTT CCTGTGAGGC   120

TTTTACCTGA CACCCGCCGC CTTTCCCCGG CACTGGCTGG GAGGGCGCCC TGCAAAGTTG   180

GGAACGCGGA GCCCCGGACC CGCTCCCGCC GCCTCCGGCT CGCCCAGGGG GGGTCGCCGG   240

GAGGAGCCCG GGGGAGAGGG ACCAGGAGGG GCCCGCGGCC TCGCAGGGGC GCCCGCGCCC   300

CCACCCCTGC CCCCGCCAGC GGACCGGTCC CCCACCCCCG GTCCTTCCAC C ATG CAC   357
                                                        Met His
                                                          1

TTG CTG GGC TTC TTC TCT GTG GCG TGT TCT CTG CTC GCC GCT GCG CTG   405
Leu Leu Gly Phe Phe Ser Val Ala Cys Ser Leu Leu Ala Ala Ala Leu
        5   10   15

CTC CCG GGT CCT CGC GAG GCG CCC GCC GCC GCC GCC GCC TTC GAG TCC   453
Leu Pro Gly Pro Arg Glu Ala Pro Ala Ala Ala Ala Ala Phe Glu Ser
        20   25   30

GGA CTC GAC CTC TCG GAC GCG GAG CCC GAC GCG GGC GAG GCC ACG GCT   501
Gly Leu Asp Leu Ser Asp Ala Glu Pro Asp Ala Gly Glu Ala Thr Ala
35   40   45   50

TAT GCA AGC AAA GAT CTG GAG GAG CAG TTA CGG TCT GTG TCC AGT GTA   549
Tyr Ala Ser Lys Asp Leu Glu Glu Gln Leu Arg Ser Val Ser Ser Val
                    55   60   65

```
GAT GAA CTC ATG ACT GTA CTC TAC CCA GAA TAT TGG AAA ATG TAC AAG        597
Asp Glu Leu Met Thr Val Leu Tyr Pro Glu Tyr Trp Lys Met Tyr Lys
            70                  75                  80

TGT CAG CTA AGG AAA GGA GGC TGG CAA CAT AAC AGA GAA CAG GCC AAC        645
Cys Gln Leu Arg Lys Gly Gly Trp Gln His Asn Arg Glu Gln Ala Asn
            85                  90                  95

CTC AAC TCA AGG ACA GAA GAG ACT ATA AAA TTT GCT GCA GCA CAT TAT        693
Leu Asn Ser Arg Thr Glu Glu Thr Ile Lys Phe Ala Ala Ala His Tyr
        100                 105                 110

AAT ACA GAG ATC TTG AAA AGT ATT GAT AAT GAG TGG AGA AAG ACT CAA        741
Asn Thr Glu Ile Leu Lys Ser Ile Asp Asn Glu Trp Arg Lys Thr Gln
115                 120                 125                 130

TGC ATG CCA CGG GAG GTG TGT ATA GAT GTG GGG AAG GAG TTT GGA GTC        789
Cys Met Pro Arg Glu Val Cys Ile Asp Val Gly Lys Glu Phe Gly Val
                135                 140                 145

GCG ACA AAC ACC TTC TTT AAA CCT CCA TGT GTG TCC GTC TAC AGA TGT        837
Ala Thr Asn Thr Phe Phe Lys Pro Pro Cys Val Ser Val Tyr Arg Cys
            150                 155                 160

GGG GGT TGC TGC AAT AGT GAG GGG CTG CAG TGC ATG AAC ACC AGC ACG        885
Gly Gly Cys Cys Asn Ser Glu Gly Leu Gln Cys Met Asn Thr Ser Thr
            165                 170                 175

AGC TAC CTC AGC AAG ACG TTA TTT GAA ATT ACA GTG CCT CTC TCT CAA        933
Ser Tyr Leu Ser Lys Thr Leu Phe Glu Ile Thr Val Pro Leu Ser Gln
        180                 185                 190

GGC CCC AAA CCA GTA ACA ATC AGT TTT GCC AAT CAC ACT TCC TGC CGA        981
Gly Pro Lys Pro Val Thr Ile Ser Phe Ala Asn His Thr Ser Cys Arg
195                 200                 205                 210

TGC ATG TCT AAA CTG GAT GTT TAC AGA CAA GTT CAT TCC ATT ATT AGA        1029
Cys Met Ser Lys Leu Asp Val Tyr Arg Gln Val His Ser Ile Ile Arg
                215                 220                 225

CGT TCC CTG CCA GCA ACA CTA CCA CAG TGT CAG GCA GCG AAC AAG ACC        1077
Arg Ser Leu Pro Ala Thr Leu Pro Gln Cys Gln Ala Ala Asn Lys Thr
            230                 235                 240

TGC CCC ACC AAT TAC ATG TGG AAT AAT CAC ATC TGC AGA TGC CTG GCT        1125
Cys Pro Thr Asn Tyr Met Trp Asn Asn His Ile Cys Arg Cys Leu Ala
            245                 250                 255

CAG GAA GAT TTT ATG TTT TCC TCG GAT GCT GGA GAT GAC TCA ACA GAT        1173
Gln Glu Asp Phe Met Phe Ser Ser Asp Ala Gly Asp Asp Ser Thr Asp
        260                 265                 270

GGA TTC CAT GAC ATC TGT GGA CCA AAC AAG GAG CTG GAT GAA GAG ACC        1221
Gly Phe His Asp Ile Cys Gly Pro Asn Lys Glu Leu Asp Glu Glu Thr
275                 280                 285                 290

TGT CAG TGT GTC TGC AGA GCG GGG CTT CGG CCT GCC AGC TGT GGA CCC        1269
Cys Gln Cys Val Cys Arg Ala Gly Leu Arg Pro Ala Ser Cys Gly Pro
                295                 300                 305

CAC AAA GAA CTA GAC AGA AAC TCA TGC CAG TGT GTC TGT AAA AAC AAA        1317
His Lys Glu Leu Asp Arg Asn Ser Cys Gln Cys Val Cys Lys Asn Lys
            310                 315                 320

CTC TTC CCC AGC CAA TGT GGG GCC AAC CGA GAA TTT GAT GAA AAC ACA        1365
Leu Phe Pro Ser Gln Cys Gly Ala Asn Arg Glu Phe Asp Glu Asn Thr
        325                 330                 335
```

```
TGC CAG TGT GTA TGT AAA AGA ACC TGC CCC AGA AAT CAA CCC CTA AAT      1413
Cys Gln Cys Val Cys Lys Arg Thr Cys Pro Arg Asn Gln Pro Leu Asn
    340                     345                 350

CCT GGA AAA TGT GCC TGT GAA TGT ACA GAA AGT CCA CAG AAA TGC TTG      1461
Pro Gly Lys Cys Ala Cys Glu Cys Thr Glu Ser Pro Gln Lys Cys Leu
355                     360             365                 370

TTA AAA GGA AAG AAG TTC CAC CAC CAA ACA TGC AGC TGT TAC AGA CGG      1509
Leu Lys Gly Lys Lys Phe His His Gln Thr Cys Ser Cys Tyr Arg Arg
                375                 380                 385

CCA TGT ACG AAC CGC CAG AAG GCT TGT GAG CCA GGA TTT TCA TAT AGT      1557
Pro Cys Thr Asn Arg Gln Lys Ala Cys Glu Pro Gly Phe Ser Tyr Ser
            390                 395                 400

GAA GAA GTG TGT CGT TGT GTC CCT TCA TAT TGG AAA AGA CCA CAA ATG      1605
Glu Glu Val Cys Arg Cys Val Pro Ser Tyr Trp Lys Arg Pro Gln Met
        405                 410                 415

AGC TAAGATTGTA CTGTTTTCCA GTTCATCGAT TTTCTATTAT GGAAAACTGT           1658
Ser

GTTGCCACAG TAGAACTGTC TGTGAACAGA GAGACCCTTG TGGGTCCATG CTAACAAAGA    1718

CAAAAGTCTG TCTTTCCTGA ACCATGTGGA TAACTTTACA GAAATGGACT GGAGCTCATC    1778

TGCAAAAGGC CTCTTGTAAA GACTGGTTTT CTGCCAATGA CCAAACAGCC AAGATTTTCC    1838

TCTTGTGATT TCTTTAAAAG AATGACTATA TAATTTATTT CCACTAAAAA TATTGTTTCT    1898

GCATTCATTT TTATAGCAAC AACAATTGGT AAAACTCACT GTGATCAATA TTTTTATATC    1958

ATGCAAAATA TGTTTAAAAT AAAATGAAAA TTGTATTAT                           1997
```

(2) INFORMATION FOR SEQ ID NO:33:

      (i) SEQUENCE CHARACTERISTICS:
         (A) LENGTH: 419 amino acids
         (B) TYPE: amino acid
         (D) TOPOLOGY: linear

     (ii) MOLECULE TYPE: protein

     (xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

```
Met His Leu Leu Gly Phe Phe Ser Val Ala Cys Ser Leu Leu Ala Ala
1               5               10                  15

Ala Leu Leu Pro Gly Pro Arg Glu Ala Pro Ala Ala Ala Ala Ala Phe
            20              25              30

Glu Ser Gly Leu Asp Leu Ser Asp Ala Glu Pro Asp Ala Gly Glu Ala
            35              40              45

Thr Ala Tyr Ala Ser Lys Asp Leu Glu Glu Gln Leu Arg Ser Val Ser
    50              55              60

Ser Val Asp Glu Leu Met Thr Val Leu Tyr Pro Glu Tyr Trp Lys Met
65              70              75              80

Tyr Lys Cys Gln Leu Arg Lys Gly Gly Trp Gln His Asn Arg Glu Gln
                85              90                  95

Ala Asn Leu Asn Ser Arg Thr Glu Glu Thr Ile Lys Phe Ala Ala Ala
            100             105             110

His Tyr Asn Thr Glu Ile Leu Lys Ser Ile Asp Asn Glu Trp Arg Lys
            115             120             125
```

```
Thr Gln Cys Met Pro Arg Glu Val Cys Ile Asp Val Gly Lys Glu Phe
    130             135             140

Gly Val Ala Thr Asn Thr Phe Phe Lys Pro Pro Cys Val Ser Val Tyr
145             150             155                 160

Arg Cys Gly Gly Cys Cys Asn Ser Glu Gly Leu Gln Cys Met Asn Thr
            165             170             175

Ser Thr Ser Tyr Leu Ser Lys Thr Leu Phe Glu Ile Thr Val Pro Leu
            180             185             190

Ser Gln Gly Pro Lys Pro Val Thr Ile Ser Phe Ala Asn His Thr Ser
        195             200             205

Cys Arg Cys Met Ser Lys Leu Asp Val Tyr Arg Gln Val His Ser Ile
    210             215             220

Ile Arg Arg Ser Leu Pro Ala Thr Leu Pro Gln Cys Gln Ala Ala Asn
225             230             235             240

Lys Thr Cys Pro Thr Asn Tyr Met Trp Asn Asn His Ile Cys Arg Cys
            245             250             255

Leu Ala Gln Glu Asp Phe Met Phe Ser Ser Asp Ala Gly Asp Asp Ser
            260             265             270

Thr Asp Gly Phe His Asp Ile Cys Gly Pro Asn Lys Glu Leu Asp Glu
    275             280             285

Glu Thr Cys Gln Cys Val Cys Arg Ala Gly Leu Arg Pro Ala Ser Cys
    290             295             300

Gly Pro His Lys Glu Leu Asp Arg Asn Ser Cys Gln Cys Val Cys Lys
305             310             315             320

Asn Lys Leu Phe Pro Ser Gln Cys Gly Ala Asn Arg Glu Phe Asp Glu
            325             330             335

Asn Thr Cys Gln Cys Val Cys Lys Arg Thr Cys Pro Arg Asn Gln Pro
            340             345             350

Leu Asn Pro Gly Lys Cys Ala Cys Glu Cys Thr Glu Ser Pro Gln Lys
        355             360             365

Cys Leu Leu Lys Gly Lys Lys Phe His His Gln Thr Cys Ser Cys Tyr
    370             375             380

Arg Arg Pro Cys Thr Asn Arg Gln Lys Ala Cys Glu Pro Gly Phe Ser
385             390             395             400

Tyr Ser Glu Glu Val Cys Arg Cys Val Pro Ser Tyr Trp Lys Arg Pro
            405             410             415

Gln Met Ser
```

(2) INFORMATION FOR SEQ ID NO:34:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 22 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: cDNA

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

TGAGTGATTTGTAGCTGCTGTG   22

(2) INFORMATION FOR SEQ ID NO:35:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 22 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: cDNA

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:

TATTGCAGCAACCCCCACATCT   22

(2) INFORMATION FOR SEQ ID NO:36:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 4416 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: cDNA

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:

```
CCACGCGCAG CGGCCGGAGA TGCAGCGGGG CGCCGCGCTG TGCCTGCGAC TGTGGCTCTG      60

CCTGGGACTC CTGGACGGCC TGGTGAGTGG CTACTCCATG ACCCCCCCGA CCTTGAACAT     120

CACGGAGGAG TCACACGTCA TCGACACCGG TGACAGCCTG TCCATCTCCT GCAGGGGACA     180

GCACCCCCTC GAGTGGGCTT GGCCAGGAGC TCAGGAGGCG CCAGCCACCG GAGACAAGGA     240

CAGCGAGGAC ACGGGGGTGG TGCGAGACTG CGAGGGCACA GACGCCAGGC CCTACTGCAA     300

GGTGTTGCTG CTGCACGAGG TACATGCCAA CGACACAGGC AGCTACGTCT GCTACTACAA     360

GTACATCAAG GCACGCATCG AGGGCACCAC GGCCGCCAGC TCCTACGTGT TCGTGAGAGA     420

CTTTGAGCAG CCATTCATCA ACAAGCCTGA CACGCTCTTG GTCAACAGGA AGGACGCCAT     480

GTGGGTGCCC TGTCTGGTGT CCATCCCCGG CCTCAATGTC ACGCTGCGCT CGCAAAGCTC     540

GGTGCTGTGG CCAGACGGGC AGGAGGTGGT GTGGGATGAC CGGCGGGGCA TGCTCGTGTC     600

CACGCCACTG CTGCACGATG CCCTGTACCT GCAGTGCGAG ACCACCTGGG GAGACCAGGA     660

CTTCCTTTCC AACCCCTTCC TGGTGCACAT CACAGGCAAC GAGCTCTATG ACATCCAGCT     720

GTTGCCCAGG AAGTCGCTGG AGCTGCTGGT AGGGGAGAAG CTGGTCCTGA ACTGCACCGT     780

GTGGGCTGAG TTTAACTCAG GTGTCACCTT TGACTGGGAC TACCCAGGGA AGCAGGCAGA     840

GCGGGGTAAG TGGGTGCCCG AGCGACGCTC CCAGCAGACC CACACAGAAC TCTCCAGCAT     900

CCTGACCATC CACAACGTCA GCCAGCACGA CCTGGGCTCG TATGTGTGCA AGGCCAACAA     960

CGGCATCCAG CGATTTCGGG AGAGCACCGA GGTCATTGTG CATGAAAATC CCTTCATCAG    1020

CGTCGAGTGG CTCAAAGGAC CCATCCTGGA GGCCACGGCA GGAGACGAGC TGGTGAAGCT    1080

GCCCGTGAAG CTGGCAGCGT ACCCCCCGCC CGAGTTCCAG TGGTACAAGG ATGGAAAGGC    1140

ACTGTCCGGG CGCCACAGTC CACATGCCCT GGTGCTCAAG GAGGTGACAG AGGCCAGCAC    1200

AGGCACCTAC ACCCTCGCCC TGTGGAACTC CGCTGCTGGC CTGAGGCGCA ACATCAGCCT    1260

GGAGCTGGTG GTGAATGTGC CCCCCCAGAT ACATGAGAAG GAGGCCTCCT CCCCCAGCAT    1320
```

CTACTCGCGT CACAGCCGCC AGGCCCTCAC CTGCACGGCC TACGGGGTGC CCCTGCCTCT   1380

CAGCATCCAG TGGCACTGGC GGCCCTGGAC ACCCTGCAAG ATGTTTGCCC AGCGTAGTCT   1440

CCGGCGGCGG CAGCAGCAAG ACCTCATGCC ACAGTGCCGT GACTGGAGGG CGGTGACCAC   1500

GCAGGATGCC GTGAACCCCA TCGAGAGCCT GGACACCTGG ACCGAGTTTG TGGAGGGAAA   1560

GAATAAGACT GTGAGCAAGC TGGTGATCCA GAATGCCAAC GTGTCTGCCA TGTACAAGTG   1620

TGTGGTCTCC AACAAGGTGG GCCAGGATGA GCGGCTCATC TACTTCTATG TGACCACCAT   1680

CCCCGACGGC TTCACCATCG AATCCAAGCC ATCCGAGGAG CTACTAGAGG GCCAGCCGGT   1740

GCTCCTGAGC TGCCAAGCCG ACAGCTACAA GTACGAGCAT CTGCGCTGGT ACCGCCTCAA   1800

CCTGTCCACG CTGCACGATG CGCACGGGAA CCCGCTTCTG CTCGACTGCA AGAACGTGCA   1860

TCTGTTCGCC ACCCCTCTGG CCGCCAGCCT GGAGGAGGTG GCACCTGGGG CGCGCCACGC   1920

CACGCTCAGC CTGAGTATCC CCGCGTCGC GCCCGAGCAC GAGGGCCACT ATGTGTGCGA   1980

AGTGCAAGAC CGGCGCAGCC ATGACAAGCA CTGCCACAAG AAGTACCTGT CGGTGCAGGC   2040

CCTGGAAGCC CCTCGGCTCA CGCAGAACTT GACCGACCTC CTGGTGAACG TGAGCGACTC   2100

GCTGGAGATG CAGTGCTTGG TGGCCGGAGC GCACGCGCCC AGCATCGTGT GGTACAAAGA   2160

CGAGAGGCTG CTGGAGGAAA AGTCTGGAGT CGACTTGGCG GACTCCAACC AGAAGCTGAG   2220

CATCCAGCGC GTGCGCGAGG AGGATGCGGG ACGCTATCTG TGCAGCGTGT GCAACGCCAA   2280

GGGCTGCGTC AACTCCTCCG CCAGCGTGGC CGTGGAAGGC TCCGAGGATA AGGGCAGCAT   2340

GGAGATCGTG ATCCTTGTCG GTACCGGCGT CATCGCTGTC TTCTTCTGGG TCCTCCTCCT   2400

CCTCATCTTC TGTAACATGA GGAGGCCGGC CCACGCAGAC ATCAAGACGG CTACCTGTC   2460

CATCATCATG GACCCCGGGG AGGTGCCTCT GGAGGAGCAA TGCGAATACC TGTCCTACGA   2520

TGCCAGCCAG TGGGAATTCC CCCGAGAGCG GCTGCACCTG GGGAGAGTGC TCGGCTACGG   2580

CGCCTTCGGG AAGGTGGTGG AAGCCTCCGC TTTCGGCATC CACAAGGGCA GCAGCTGTGA   2640

CACCGTGGCC GTGAAAATGC TGAAAGAGGG CGCCACGGCC AGCGAGCACC GCGCGCTGAT   2700

GTCGGAGCTC AAGATCCTCA TTCACATCGG CAACCACCTC AACGTGGTCA ACCTCCTCGG   2760

GGCGTGCACC AAGCCGCAGG GCCCCCTCAT GGTGATCGTG GAGTTCTGCA AGTACGGCAA   2820

CCTCTCCAAC TTCCTGCGCG CCAAGCGGGA CGCCTTCAGC CCCTGCGCGG AGAAGTCTCC   2880

CGAGCAGCGC GGACGCTTCC GCGCCATGGT GGAGCTCGCC AGGCTGGATC GGAGGCGGCC   2940

GGGGAGCAGC GACAGGGTCC TCTTCGCGCG GTTCTCGAAG ACCGAGGGCG GAGCGAGGCG   3000

GGCTTCTCCA GACCAAGAAG CTGAGGACCT GTGGCTGAGC CCGCTGACCA TGGAAGATCT   3060

TGTCTGCTAC AGCTTCCAGG TGGCCAGAGG GATGGAGTTC CTGGCTTCCC GAAAGTGCAT   3120

CCACAGAGAC CTGGCTGCTC GGAACATTCT GCTGTCGGAA AGCGACGTGG TGAAGATCTG   3180

TGACTTTGGC CTTGCCCGGG ACATCTACAA AGACCCTGAC TACGTCCGCA AGGGCAGTGC   3240

CCGGCTGCCC CTGAAGTGGA TGGCCCCTGA AAGCATCTTC GACAAGGTGT ACACCACGCA   3300

GAGTGACGTG TGGTCCTTTG GGGTGCTTCT CTGGGAGATC TTCTCTCTGG GGGCCTCCCC   3360

GTACCCTGGG GTGCAGATCA ATGAGGAGTT CTGCCAGCGG CTGAGAGACG GCACAAGGAT   3420

GAGGGCCCCG GAGCTGGCCA CTCCCGCCAT ACGCCGCATC ATGCTGAACT GCTGGTCCGG   3480

AGACCCCAAG GCGAGACCTG CATTCTCGGA GCTGGTGGAG ATCCTGGGGG ACCTGCTCCA   3540

GGGCAGGGGC CTGCAAGAGG AAGAGGAGGT CTGCATGGCC CCGCGCAGCT CTCAGAGCTC   3600

AGAAGAGGGC AGCTTCTCGC AGGTGTCCAC CATGGCCCTA CACATCGCCC AGGCTGACGC   3660

TGAGGACAGC CCGCCAAGCC TGCAGCGCCA CAGCCTGGCC GCCAGGTATT ACAACTGGGT   3720

GTCCTTTCCC GGGTGCCTGG CCAGAGGGGC TGAGACCCGT GGTTCCTCCA GGATGAAGAC   3780

ATTTGAGGAA TTCCCCATGA CCCCAACGAC CTACAAAGGC TCTGTGGACA ACCAGACAGA   3840

CAGTGGGATG GTGCTGGCCT CGGAGGAGTT TGAGCAGATA GAGAGCAGGC ATAGACAAGA   3900

AAGCGGCTTC AGGTAGCTGA AGCAGAGAGA GAGAAGGCAG CATACGTCAG CATTTTCTTC   3960

TCTGCACTTA TAAGAAAGAT CAAAGACTTT AAGACTTTCG CTATTTCTTC TACTGCTATC   4020

TACTACAAAC TTCAAAGAGG AACCAGGAGG ACAAGAGGAG CATGAAAGTG ACAAGGAGT   4080

GTGACCACTG AAGCACCACA GGGAAGGGGT TAGGCCTCCG GATGACTGCG GGCAGGCCTG   4140

GATAATATCC AGCCTCCCAC AAGAAGCTGG TGGAGCAGAG TGTTCCCTGA CTCCTCCAAG   4200

GAAAGGGAGA CGCCCTTTCA TGGTCTGCTG AGTAACAGGT GCNTTCCCAG ACACTGGCGT   4260

TACTGCTTGA CCAAAGAGCC CTCAAGCGGC CCTTATGCCA GCGTGACAGA GGGCTCACCT   4320

CTTGCCTTCT AGGTCACTTC TCACACAATG TCCCTTCAGC ACCTGACCCT GTGCCCGCCA   4380

GTTATTCCTT GGTAATATGA GTAATACATC AAAGAG   4416

(2) INFORMATION FOR SEQ ID NO:37:

         (i) SEQUENCE CHARACTERISTICS:
               (A) LENGTH: 4273 base pairs
               (B) TYPE: nucleic acid
               (C) STRANDEDNESS: single
               (D) TOPOLOGY: linear

         (ii) MOLECULE TYPE: cDNA

         (xi) SEQUENCE DESCRIPTION: SEQ ID NO:37:

AAGCTTATCG ATTTCGAACC CGGGGGTACC GAATTCCTCG AGTCTAGAGG AGCATGCCTG   60

CAGGTCGACC GGGCTCGATC CCCTCGCGAG TTGGTTCAGC TGCTGCCTGA GGCTGGACGA   120

CCTCGCGGAG TTCTACCGGC AGTGCAAATC CGTCGGCATC CAGGAAACCA GCAGCGGCTA   180

TCCGCGCATC CATGCCCCCG AACTGCAGGA GTGGGGAGGC ACGATGGCCG CTTTGGTCCC   240

GGATCTTTGT GAAGGAACCT TACTTCTGTG GTGTGACATA ATTGGACAAA CTACCTACAG   300

AGATTTAAAG CTCTAAGGTA AATATAAAAT TTTTAAGTGT ATAATGTGTT AAACTACTGA   360

TTCTAATTGT TTGTGTATTT TAGATTCCAA CCTATGGAAC TGATGAATGG GAGCAGTGGT   420

GGAATGCCTT TAATGAGGAA AACCTGTTTT GCTCAGAAGA AATGCCATCT AGTGATGATG   480

AGGCTACTGC TGACTCTCAA CATTCTACTC CTCCAAAAAA GAAGAGAAAG GTAGAAGACC   540

CCAAGGACTT TCCTTCAGAA TTGCTAAGTT TTTTGAGTCA TGCTGTGTTT AGTAATAGAA   600

CTCTTGCTTG CTTTGCTATT TACACCACAA AGGAAAAAGC TGCACTGCTA TACAAGAAAA   660

TTATGGAAAA ATATTCTGTA ACCTTTATAA GTAGGCATAA CAGTTATAAT CATAACATAC   720

```
TGTTTTTTCT TACTCCACAC AGGCATAGAG TGTCTGCTAT TAATAACTAT GCTCAAAAAT    780
TGTGTACCTT TAGCTTTTTA ATTTGTAAAG GGGTTAATAA GGAATATTTG ATGTATAGTG    840
CCTTGACTAG AGATCATAAT CAGCCATACC ACATTTGTAG AGGTTTTACT TGCTTTAAAA    900
AACCTCCCAC ACCTCCCCCT GAACCTGAAA CATAAAATGA ATGCAATTGT TGTTGTTAAC    960
TTGTTTATTG CAGCTTATAA TGGTTACAAA TAAAGCAATA GCATCACAAA TTTCACAAAT   1020
AAAGCATTTT TTTCACTGCA TTCTAGTTGT GGTTTGTCCA AACTCATCAA TGTATCTTAT   1080
CATGTCTGGA TCTGCCGGTC TCCCTATAGT GAGTCGTATT AATTTCGATA AGCCAGGTTA   1140
ACCTGCATTA ATGAATCGGC CAACGCGCGG GGAGAGGCGG TTTGCGTATT GGGCGCTCTT   1200
CCGCTTCCTC GCTCACTGAC TCGCTGCGCT CGGTCGTTCG GCTGCGGCGA GCGGTATCAG   1260
CTCACTCAAA GGCGGTAATA CGGTTATCCA CAGAATCAGG GGATAACGCA GGAAAGAACA   1320
TGTGAGCAAA AGGCCAGCAA AAGGCCAGGA ACCGTAAAAA GGACGCGTTG CTGGCGTTTT   1380
TCCATAGGCT CCGCCCCCCT GACGAGCATC ACAAAAATCG ACGCTCAAGT CAGAGGTGGC   1440
GAAACCCGAC AGGACTATAA AGATACCAGG CGTTTCCCCC TGGAAGCTCC CTCGTGCGCT   1500
CTCCTGTTCC GACCCTGCCG CTTACCGGAT ACCTGTCCGC CTTTCTCCCT TCGGGAAGCG   1560
TGGCGCTTTC TCAATGCTCA CGCTGTAGGT ATCTCAGTTC GGTGTAGGTC GTTCGCTCCA   1620
AGCTGGGCTG TGTGCACGAA CCCCCCGTTC AGCCCGACCG CTGCGCCTTA TCCGGTAACT   1680
ATCGTCTTGA GTCCAACCCG GTAAGACACG ACTTATCGCC ACTGGCAGCA GCCACTGGTA   1740
ACAGGATTAG CAGAGCGAGG TATGTAGGCG GTGCTACAGA GTTCTTGAAG TGGTGGCCTA   1800
ACTACGGCTA CACTAGAAGG ACAGTATTTG GTATCTGCGC TCTGCTGAAG CCAGTTACCT   1860
TCGGAAAAAG AGTTGGTAGC TCTTGATCCG GCAAACAAAC CACCGCTGGT AGCGGTGGTT   1920
TTTTTGTTTG CAAGCAGCAG ATTACGCGCA GAAAAAAAGG ATCTCAAGAA GATCCTTTGA   1980
TCTTTTCTAC GGGGTCTGAC GCTCAGTGGA ACGAAAACTC ACGTTAAGGG ATTTTGGTCA   2040
TGAGATTATC AAAAAGGATC TTCACCTAGA TCCTTTTAAA TTAAAAATGA AGTTTTAAAT   2100
CAATCTAAAG TATATATGAG TAAACTTGGT CTGACAGTTA CCAATGCTTA ATCAGTGAGG   2160
CACCTATCTC AGCGATCTGT CTATTTCGTT CATCCATAGT TGCCTGACTC CCCGTCGTGT   2220
AGATAACTAC GATACGGGAG GGCTTACCAT CTGGCCCCAG TGCTGCAATG ATACCGCGAG   2280
ACCCACGCTC ACCGGCTCCA GATTTATCAG CAATAAACCA GCCAGCCGGA AGGGCCGAGC   2340
GCAGAAGTGG TCCTGCAACT TTATCCGCCT CCATCCAGTC TATTAATTGT TGCCGGGAAG   2400
CTAGAGTAAG TAGTTCGCCA GTTAATAGTT TGCGCAACGT TGTTGCCATT GCTACAGGCA   2460
TCGTGGTGTC ACGCTCGTCG TTTGGTATGG CTTCATTCAG CTCCGGTTCC CAACGATCAA   2520
GGCGAGTTAC ATGATCCCCC ATGTTGTGCA AAAAAGCGGT TAGCTCCTTC GGTCCTCCGA   2580
TCGTTGTCAG AAGTAAGTTG GCCGCAGTGT TATCACTCAT GGTTATGGCA GCACTGCATA   2640
ATTCTCTTAC TGTCATGCCA TCCGTAAGAT GCTTTTCTGT GACTGGTGAG TACTCAACCA   2700
AGTCATTCTG AGAATAGTGT ATGCGGCGAC CGAGTTGCTC TTGCCCGGCG TCAATACGGG   2760
ATAATACCGC GCCACATAGC AGAACTTTAA AAGTGCTCAT CATTGGAAAA CGTTCTTCGG   2820
```

```
GGCGAAAACT CTCAAGGATC TTACCGCTGT TGAGATCCAG TTCGATGTAA CCCACTCGTG    2880

CACCCAACTG ATCTTCAGCA TCTTTTACTT TCACCAGCGT TTCTGGGTGA GCAAAAACAG    2940

GAAGGCAAAA TGCCGCAAAA AAGGGAATAA GGGCGACACG GAAATGTTGA ATACTCATAC    3000

TCTTCCTTTT TCAATATTAT TGAAGCATTT ATCAGGGTTA TTGTCTCATG AGCGGATACA    3060

TATTTGAATG TATTTAGAAA AATAAACAAA TAGGGGTTCC GCGCACATTT CCCCGAAAAG    3120

TGCCACCTGA CGTCTAAGAA ACCATTATTA TCATGACATT AACCTATAAA AATAGGCGTA    3180

TCACGAGGCC CTTTCGTCTC GCGCGTTTCG GTGATGACGG TGAAAACCTC TGACACATGC    3240

AGCTCCCGGA GACGGTCACA GCTTGTCTGT AAGCGGATGC CGGGAGCAGA CAAGCCCGTC    3300

AGGGCGCGTC AGCGGGTGTT GGCGGGTGTC GGGGCTGGCT TAACTATGCG GCATCAGAGC    3360

AGATTGTACT GAGAGTGCAC CATATGGACA TATTGTCGTT AGAACGCGGC TACAATTAAT    3420

ACATAACCTT ATGTATCATA CACATACGAT TTAGGTGACA CTATAGAACT CGAGCAGAGC    3480

TTCCAAATTG AGAGAGAGGC TTAATCAGAG ACAGAAACTG TTTGAGTCAA CTCAAGGATG    3540

GTTTGAGGGA CTGTTTAACA GATCCCCTTG GTTTACCACC TTGATATCTA CCATTATGGG    3600

ACCCCTCATT GTACTCCTAA TGATTTTGCT CTTCGGACCC TGCATTCTTA ATCGATTAGT    3660

CCAATTTGTT AAAGACAGGA TATCAGTGGT CCAGGCTCTA GTTTTGACTC AACAATATCA    3720

CCAGCTGAAG CCTATAGAGT ACGAGCCATA GATAAAATAA AAGATTTTAT TTAGTCTCCA    3780

GAAAAAGGGG GGAATGAAAG ACCCCACCTG TAGGTTTGGC AAGCTAGCTT AAGTAACGCC    3840

ATTTTGCAAG GCATGGAAAA ATACATAACT GAGAATAGAG AAGTTCAGAT CAAGGTCAGG    3900

AACAGATGGA ACAGCTGAAT ATGGGCCAAA CAGGATATCT GTGGTAAGCA GTTCCTGCCC    3960

CGGCTCAGGG CCAAGAACAG ATGGAACAGC TGAATATGGG CCAAACAGGA TATCTGTGGT    4020

AAGCAGTTCC TGCCCCGGCT CAGGGCCAAG AACAGATGGT CCCCAGATGC GGTCCAGCCC    4080

TCAGCAGTTT CTAGAGAACC ATCAGATGTT TCCAGGGTGC CCCAAGGACC TGAAATGACC    4140

CTGTGCCTTA TTTGAACTAA CCAATCAGTT CGCTTCTCGC TTCTGTTCGC GCGCTTCTGC    4200

TCCCCGAGCT CAATAAAAGA GCCCACAACC CCTCACTCGG GGCGCCAGTC CTCCGATTGA    4260

CTGAGTCGCC CGG                                                      4273
```

(2) INFORMATION FOR SEQ ID NO:38:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 216 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: cDNA

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:38:

```
CAAGAAAGCG GCTTCAGCTG TAAAGGACCT GGCCAGAATG TGGCTGTGAC CAGGGCACAC     60

CCTGACTCCC AAGGGAGGCG GCGGCGGCCT GAGCGGGGGG CCCGAGGAGG CCAGGTGTTT    120

TACAACAGCG AGTATGGGGA GCTGTCGGAG CCAAGCGAGG AGGACCACTG CTCCCCGTCT    180

GCCCGCGTGA CTTTCTTCAC AGACAACAGC TACTAA                             216
```

(2) INFORMATION FOR SEQ ID NO:39:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 17 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:39:

```
Glu Glu Thr Ile Lys Phe Ala Ala Ala His Tyr Asn Thr Glu Ile Leu
1               5                   10                  15
```

Lys

(2) INFORMATION FOR SEQ ID NO:40:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 1836 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: cDNA

    (ix) FEATURE:
        (A) NAME/KEY: CDS
        (B) LOCATION: 168..1412

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:40:

```
GCGGCCGCGT CGACGCAAAA GTTGCGAGCC GCCGAGTCCC GGGAGACGCT CGCCCAGGGG        60

GGTCCCCGGG AGGAAACCAC GGGACAGGGA CCAGGAGAGG ACCTCAGCCT CACGCCCCAG       120

CCTGCGCCAG CCAACGGACC GGCCTCCCTG CTCCCGGTCC ATCCACC ATG CAC TTG        176
                                                    Met His Leu
                                                     1

CTG TGC TTC TTG TCT CTG GCG TGT TCC CTG CTC GCC GCT GCG CTG ATC        224
Leu Cys Phe Leu Ser Leu Ala Cys Ser Leu Leu Ala Ala Ala Leu Ile
        5                   10                  15

CCC AGT CCG CGC GAG GCG CCC GCC ACC GTC GCC GCC TTC GAG TCG GGA        272
Pro Ser Pro Arg Glu Ala Pro Ala Thr Val Ala Ala Phe Glu Ser Gly
 20                  25                  30                  35

CTG GGC TTC TCG GAA GCG GAG CCC GAC GGG GGC GAG GTC AAG GCT TTT        320
Leu Gly Phe Ser Glu Ala Glu Pro Asp Gly Gly Glu Val Lys Ala Phe
                40                  45                  50

GAA GGC AAA GAC CTG GAG GAG CAG TTG CGG TCT GTG TCC AGC GTA GAT        368
Glu Gly Lys Asp Leu Glu Glu Gln Leu Arg Ser Val Ser Ser Val Asp
                55                  60                  65

GAG CTG ATG TCT GTC CTG TAC CCA GAC TAC TGG AAA ATG TAC AAG TGC        416
Glu Leu Met Ser Val Leu Tyr Pro Asp Tyr Trp Lys Met Tyr Lys Cys
                70                  75                  80

CAG CTG CGG AAA GGC GGC TGG CAG CAG CCC ACC CTC AAT ACC AGG ACA        464
Gln Leu Arg Lys Gly Gly Trp Gln Gln Pro Thr Leu Asn Thr Arg Thr
        85                  90                  95

GGG GAC AGT GTA AAA TTT GCT GCT GCA CAT TAT AAC ACA GAG ATC CTG        512
Gly Asp Ser Val Lys Phe Ala Ala Ala His Tyr Asn Thr Glu Ile Leu
100                 105                 110                 115
```

```
AAA AGT ATT GAT AAT GAG TGG AGA AAG ACT CAA TGC ATG CCA CGT GAG          560
Lys Ser Ile Asp Asn Glu Trp Arg Lys Thr Gln Cys Met Pro Arg Glu
            120                 125                 130

GTG TGT ATA GAT GTG GGG AAG GAG TTT GGA GCA GCC ACA AAC ACC TTC          608
Val Cys Ile Asp Val Gly Lys Glu Phe Gly Ala Ala Thr Asn Thr Phe
            135                 140                 145

TTT AAA CCT CCA TGT GTG TCC GTC TAC AGA TGT GGG GGT TGC TGC AAC          656
Phe Lys Pro Pro Cys Val Ser Val Tyr Arg Cys Gly Gly Cys Cys Asn
            150                 155                 160

AGG GAG GGG CTG CAG TGC ATG AAC ACC AGC ACA GGT TAC CTC AGC AAG          704
Arg Glu Gly Leu Gln Cys Met Asn Thr Ser Thr Gly Tyr Leu Ser Lys
            165                 170                 175

ACG TTG TTT GAA ATT ACA GTG CCT CTC TCA CAA GGC CCC AAA CCA GTC          752
Thr Leu Phe Glu Ile Thr Val Pro Leu Ser Gln Gly Pro Lys Pro Val
180                 185                 190                 195

ACA ATC AGT TTT GCC AAT CAC ACT TCC TGC CGG TGC ATG TCT AAA CTG          800
Thr Ile Ser Phe Ala Asn His Thr Ser Cys Arg Cys Met Ser Lys Leu
            200                 205                 210

GAT GTT TAC AGA CAA GTT CAT TCA ATT ATT AGA CGT TCT CTG CCA GCA          848
Asp Val Tyr Arg Gln Val His Ser Ile Ile Arg Arg Ser Leu Pro Ala
            215                 220                 225

ACA TTA CCA CAG TGT CAG GCA GCT AAC AAG ACA TGT CCA ACA AAC TAT          896
Thr Leu Pro Gln Cys Gln Ala Ala Asn Lys Thr Cys Pro Thr Asn Tyr
            230                 235                 240

GTG TGG AAT AAC TAC ATG TGC CGA TGC CTG GCT CAG CAG GAT TTT ATC          944
Val Trp Asn Asn Tyr Met Cys Arg Cys Leu Ala Gln Gln Asp Phe Ile
            245                 250                 255

TTT TAT TCA AAT GTT GAA GAT GAC TCA ACC AAT GGA TTC CAT GAT GTC          992
Phe Tyr Ser Asn Val Glu Asp Asp Ser Thr Asn Gly Phe His Asp Val
260                 265                 270                 275

TGT GGA CCC AAC AAG GAG CTG GAT GAA GAC ACC TGT CAG TGT GTC TGC          1040
Cys Gly Pro Asn Lys Glu Leu Asp Glu Asp Thr Cys Gln Cys Val Cys
            280                 285                 290

AAG GGG GGG CTT CGG CCA TCT AGT TGT GGA CCC CAC AAA GAA CTA GAT          1088
Lys Gly Gly Leu Arg Pro Ser Ser Cys Gly Pro His Lys Glu Leu Asp
            295                 300                 305

AGA GAC TCA TGT CAG TGT GTC TGT AAA AAC AAA CTT TTC CCT AAT TCA          1136
Arg Asp Ser Cys Gln Cys Val Cys Lys Asn Lys Leu Phe Pro Asn Ser
            310                 315                 320

TGT GGA GCC AAC AGG GAA TTT GAT GAG AAT ACA TGT CAG TGT GTA TGT          1184
Cys Gly Ala Asn Arg Glu Phe Asp Glu Asn Thr Cys Gln Cys Val Cys
325                 330                 335

AAA AGA ACG TGT CCA AGA AAT CAG CCC CTG AAT CCT GGG AAA TGT GCC          1232
Lys Arg Thr Cys Pro Arg Asn Gln Pro Leu Asn Pro Gly Lys Cys Ala
340                 345                 350                 355

TGT GAA TGT ACA GAA AAC ACA CAG AAG TGC TTC CTT AAA GGG AAG AAG          1280
Cys Glu Cys Thr Glu Asn Thr Gln Lys Cys Phe Leu Lys Gly Lys Lys
            360                 365                 370

TTC CAC CAT CAA ACA TGC AGT TGT TAC AGA AGA CCG TGT GCG AAT CGA          1328
Phe His His Gln Thr Cys Ser Cys Tyr Arg Arg Pro Cys Ala Asn Arg
            375                 380                 385
```

CTG AAG CAT TGT GAT CCA GGA CTG TCC TTT AGT GAA GAA GTA TGC CGC   1376
Leu Lys His Cys Asp Pro Gly Leu Ser Phe Ser Glu Glu Val Cys Arg
        390   395   400

TGT GTC CCA TCG TAT TGG AAA AGG CCA CAT CTG AAC TAAGATCATA   1422
Cys Val Pro Ser Tyr Trp Lys Arg Pro His Leu Asn
        405   410   415

CCAGTTTTCA GTCAGTCACA GTCATTTACT CTCTTGAAGA CTGTTGGAAC AGCACTTAGC   1482

ACTGTCTATG CACAGAAAGA CTCTGTGGGA CCACATGGTA ACAGAGGCCC AAGTCTGTGT   1542

TTATTGAACC ATGTGGATTA CTGCGGGAGA GGACTGGCAC TCATGTGCAA AAAAAACCTC   1602

TTCAAAGACT GGTTTCTGC CAGGGACCAG ACAGCTGAGG TTTTTCTCTT GTGATTTAAA   1662

AAAAGAATGA CTATATAATT TATTTCCACT AAAAATATTG TTCCTGCATT CATTTTTATA   1722

GCAATAACAA TTGGTAAAGC TCACTGTGAT CAGTATTTTT ATAACATGCA AAACTATGTT   1782

TAAAATAAAA TGAAAATTGT ATTATAAAAA AAAAAAAAA AAAAAAAAA GCTT   1836

(2) INFORMATION FOR SEQ ID NO:41:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 415 amino acids
        (B) TYPE: amino acid
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: protein

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:41:

Met His Leu Leu Cys Phe Leu Ser Leu Ala Cys Ser Leu Leu Ala Ala
1   5   10   15

Ala Leu Ile Pro Ser Pro Arg Glu Ala Pro Ala Thr Val Ala Ala Phe
            20   25   30

Glu Ser Gly Leu Gly Phe Ser Glu Ala Glu Pro Asp Gly Gly Glu Val
        35   40   45

Lys Ala Phe Glu Gly Lys Asp Leu Glu Glu Gln Leu Arg Ser Val Ser
    50   55   60

Ser Val Asp Glu Leu Met Ser Val Leu Tyr Pro Asp Tyr Trp Lys Met
65   70   75   80

Tyr Lys Cys Gln Leu Arg Lys Gly Gly Trp Gln Gln Pro Thr Leu Asn
                85   90   95

Thr Arg Thr Gly Asp Ser Val Lys Phe Ala Ala Ala His Tyr Asn Thr
            100   105   110

Glu Ile Leu Lys Ser Ile Asp Asn Glu Trp Arg Lys Thr Gln Cys Met
        115   120   125

Pro Arg Glu Val Cys Ile Asp Val Gly Lys Glu Phe Gly Ala Ala Thr
    130   135   140

Asn Thr Phe Phe Lys Pro Pro Cys Val Ser Val Tyr Arg Cys Gly Gly
145   150   155   160

Cys Cys Asn Arg Glu Gly Leu Gln Cys Met Asn Thr Ser Thr Gly Tyr
            165   170   175

Leu Ser Lys Thr Leu Phe Glu Ile Thr Val Pro Leu Ser Gln Gly Pro
            180   185   190

```
Lys Pro Val Thr Ile Ser Phe Ala Asn His Thr Ser Cys Arg Cys Met
        195             200             205

Ser Lys Leu Asp Val Tyr Arg Gln Val His Ser Ile Ile Arg Arg Ser
    210             215             220

Leu Pro Ala Thr Leu Pro Gln Cys Gln Ala Ala Asn Lys Thr Cys Pro
225             230             235             240

Thr Asn Tyr Val Trp Asn Asn Tyr Met Cys Arg Cys Leu Ala Gln Gln
            245             250             255

Asp Phe Ile Phe Tyr Ser Asn Val Glu Asp Asp Ser Thr Asn Gly Phe
            260             265             270

His Asp Val Cys Gly Pro Asn Lys Glu Leu Asp Glu Asp Thr Cys Gln
        275             280             285

Cys Val Cys Lys Gly Gly Leu Arg Pro Ser Ser Cys Gly Pro His Lys
        290             295             300

Glu Leu Asp Arg Asp Ser Cys Gln Cys Val Cys Lys Asn Lys Leu Phe
305             310             315             320

Pro Asn Ser Cys Gly Ala Asn Arg Glu Phe Asp Glu Asn Thr Cys Gln
            325             330             335

Cys Val Cys Lys Arg Thr Cys Pro Arg Asn Gln Pro Leu Asn Pro Gly
        340             345             350

Lys Cys Ala Cys Glu Cys Thr Glu Asn Thr Gln Lys Cys Phe Leu Lys
        355             360             365

Gly Lys Lys Phe His His Gln Thr Cys Ser Cys Tyr Arg Arg Pro Cys
370             375             380

Ala Asn Arg Leu Lys His Cys Asp Pro Gly Leu Ser Phe Ser Glu Glu
385             390             395             400

Val Cys Arg Cys Val Pro Ser Tyr Trp Lys Arg Pro His Leu Asn
            405             410             415
```

(2) INFORMATION FOR SEQ ID NO:42:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 22 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: cDNA

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:42:

TCTCTTCTGT GCTTGAGTTG AG   22

(2) INFORMATION FOR SEQ ID NO:43:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 22 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: cDNA

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:43:

TCTCTTCTGT CCCTGAGTTG AG   22

(2) INFORMATION FOR SEQ ID NO:44:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 65 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: cDNA

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:44:

TGTGCTGCAG CAAATTTTAT AGTCTCTTCT GTGGCGGCGG CGGCGGCGGG CGCCTCGCGA   60

GGACC   65

(2) INFORMATION FOR SEQ ID NO:45:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 30 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: cDNA

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:45:

CTGGCAGGGA ACTGCTAATA ATGGAATGAA   30

(2) INFORMATION FOR SEQ ID NO:46:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 84 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: cDNA

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:46:

GGGCTCCGCG TCCGAGAGGT CGAGTCCGGA CTCGTGATGG TGATGGTGAT GGGCGGCGGC   60

GGCGGCGGGC GCCTCGCGAG GACC   84

(2) INFORMATION FOR SEQ ID NO:47:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 31 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: cDNA

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:47:

GTATTATAAT GTCCTCCACC AAATTTTATA G   31

(2) INFORMATION FOR SEQ ID NO:48:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 93 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: cDNA

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:48:

GTTCGCTGCC TGACACTGTG GTAGTGTTGC TGGCGGCCGC TAGTGATGGT GATGGTGATG   60

AATAATGGAA TGAACTTGTC TGTAAACATC CAG   93

    (2) INFORMATION FOR SEQ ID NO:49:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 18 base pairs
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: single
            (D) TOPOLOGY: linear

        (ii) MOLECULE TYPE: cDNA

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:49:

CATGTACGAA CCGCCAGG   18

    (2) INFORMATION FOR SEQ ID NO:50:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 20 base pairs
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: single
            (D) TOPOLOGY: linear

        (ii) MOLECULE TYPE: cDNA

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:50:

AATGACCAGA GAGAGGCGAG   20

    (2) INFORMATION FOR SEQ ID NO:51:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 10 amino acids
            (B) TYPE: amino acid
            (C) STRANDEDNESS: single
            (D) TOPOLOGY: linear

        (ii) MOLECULE TYPE: peptide

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:51:

    Ala Val Val Met Thr Gln Thr Pro Ala Ser
    1   5   10

    (2) INFORMATION FOR SEQ ID NO:52:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 1741 base pairs
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: single
            (D) TOPOLOGY: linear

        (ii) MOLECULE TYPE: cDNA

        (ix) FEATURE:
            (A) NAME/KEY: CDS
            (B) LOCATION: 453..1706

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:52:

GCCCCCGCCG AGCGCTCCGC GCGCAGCCGC CGGGCCGGGC CGGCCCGCGG AGGGCGCGCT   60

GCGAGCGGCC ACTGGGTCCT GCTTCCCTCC TTCCTCTCCC TCCTCCTCCT CCTCCTTCTC   120

TCTGCGCTTT CCACCGCTCC CGAGCGAGCG CACGCTCGGA TGTCCGGTTT CCTGGTGGGT   180

TTTTTACCTG GCAAAGTCCG GATAACTTCG GTGAGAATTT GCAAAGAGGC TGGGAGCTCC   240

CCTGCAGGCG TCTGGGAGCT GCTGCCGCCG TCGCATCTTC TCCATCCCGC GGATTTTACT   300

GCCTTGGATA TTGCGAGGGG AGGGAGGGGG GTGAGGACAG CAAAAAGAAA GGGGTGGGGG   360

GGGGGAGAGA AAAGGAAAAG AAGGAGCCTC GGAATTGTGC CCGCATTCCT GCGCTGCCCC   420

GCGGCCCCCC TCCGCTCTGC CATCTCCGCA CA ATG CAC TTG CTG GAG ATG CTC   473
                                    Met His Leu Leu Glu Met Leu
                                     1   5

TCC CTG GGC TGC TGC CTC GCT GCT GGC GCC GTG CTC CTG GGA CCC CGG   521
Ser Leu Gly Cys Cys Leu Ala Ala Gly Ala Val Leu Leu Gly Pro Arg
            10   15   20

CAG CCG CCC GTC GCC GCC GCC TAC GAG TCC GGG CAC GGC TAC TAC GAG   569
Gln Pro Pro Val Ala Ala Ala Tyr Glu Ser Gly His Gly Tyr Tyr Glu
        25   30   35

GAG GAG CCC GGT GCC GGG GAA CCC AAG GCT CAT GCA AGC AAA GAC CTG   617
Glu Glu Pro Gly Ala Gly Glu Pro Lys Ala His Ala Ser Lys Asp Leu
 40   45   50   55

GAA GAG CAG TTG CGA TCT GTG TCC AGT GTG GAT GAA CTC ATG ACA GTA   665
Glu Glu Gln Leu Arg Ser Val Ser Ser Val Asp Glu Leu Met Thr Val
            60   65   70

CTT TAC CCA GAA TAC TGG AAA ATG TTC AAA TGT CAG TTG AGG AAA GGA   713
Leu Tyr Pro Glu Tyr Trp Lys Met Phe Lys Cys Gln Leu Arg Lys Gly
            75   80   85

GGT TGG CAA CAC AAC AGG GAA CAC TCC AGC TCT GAT ACA AGA TCA GAT   761
Gly Trp Gln His Asn Arg Glu His Ser Ser Ser Asp Thr Arg Ser Asp
            90   95   100

GAT TCA TTG AAA TTT GCC GCA GCA CAT TAT AAT GCA GAG ATC CTG AAA   809
Asp Ser Leu Lys Phe Ala Ala Ala His Tyr Asn Ala Glu Ile Leu Lys
        105   110   115

AGT ATT GAT ACT GAA TGG AGA AAA ACC CAG GGC ATG CCA CGT GAA GTG   857
Ser Ile Asp Thr Glu Trp Arg Lys Thr Gln Gly Met Pro Arg Glu Val
120   125   130   135

TGT GTG GAT TTG GGG AAA GAG TTT GGA GCA ACT ACA AAC ACC TTC TTT   905
Cys Val Asp Leu Gly Lys Glu Phe Gly Ala Thr Thr Asn Thr Phe Phe
            140   145   150

AAA CCC CCG TGT GTA TCC ATC TAC AGA TGT GGA GGT TGC TGC AAT AGT   953
Lys Pro Pro Cys Val Ser Ile Tyr Arg Cys Gly Gly Cys Cys Asn Ser
            155   160   165

GAA GGA CTC CAG TGT ATG AAT ATC AGC ACA AAT TAC ATC AGC AAG ACA   1001
Glu Gly Leu Gln Cys Met Asn Ile Ser Thr Asn Tyr Ile Ser Lys Thr
            170   175   180

TTG TTT GAG ATT ACA GTG CCT CTC TCT CAT GGC CCC AAA CCT GTA ACA   1049
Leu Phe Glu Ile Thr Val Pro Leu Ser His Gly Pro Lys Pro Val Thr
        185   190   195

GTC AGT TTT GCC AAT CAC ACG TCC TGC CGA TGC ATG TCT AAG TTG GAT   1097
Val Ser Phe Ala Asn His Thr Ser Cys Arg Cys Met Ser Lys Leu Asp
200   205   210   215

GTT TAC AGA CAA GTT CAT TCT ATC ATA AGA CGT TCC TTG CCA GCA ACA   1145
Val Tyr Arg Gln Val His Ser Ile Ile Arg Arg Ser Leu Pro Ala Thr
            220   225   230

```
CAA ACT CAG TGT CAT GTG GCA AAC AAG ACC TGT CCA AAA AAT CAT GTC      1193
Gln Thr Gln Cys His Val Ala Asn Lys Thr Cys Pro Lys Asn His Val
        235                 240                 245

TGG AAT AAT CAG ATT TGC AGA TGC TTA GCA CAG CAC GAT TTT GGT TTC      1241
Trp Asn Asn Gln Ile Cys Arg Cys Leu Ala Gln His Asp Phe Gly Phe
        250                 255                 260

TCT TCT CAC CTT GGA GAT TCT GAC ACA TCT GAA GGA TTC CAT ATT TGT      1289
Ser Ser His Leu Gly Asp Ser Asp Thr Ser Glu Gly Phe His Ile Cys
        265                 270                 275

GGG CCC AAC AAA GAG CTG GAT GAA GAA ACC TGT CAA TGC GTC TGC AAA      1337
Gly Pro Asn Lys Glu Leu Asp Glu Glu Thr Cys Gln Cys Val Cys Lys
280                 285                 290                 295

GGA GGT GTG CGG CCC ATA AGC TGT GGC CCT CAC AAA GAA CTA GAC AGG      1385
Gly Gly Val Arg Pro Ile Ser Cys Gly Pro His Lys Glu Leu Asp Arg
                300                 305                 310

GCA TCA TGT CAG TGC ATG TGC AAA AAC AAA CTG CTC CCC AGT TCC TGT      1433
Ala Ser Cys Gln Cys Met Cys Lys Asn Lys Leu Leu Pro Ser Ser Cys
        315                 320                 325

GGG CCT AAC AAA GAA TTT GAT GAA GAA AAG TGC CAG TGT GTA TGT AAA      1481
Gly Pro Asn Lys Glu Phe Asp Glu Glu Lys Cys Gln Cys Val Cys Lys
        330                 335                 340

AAG ACC TGT CCC AAA CAT CAT CCA CTA AAT CCT GCA AAA TGC ATC TGC      1529
Lys Thr Cys Pro Lys His His Pro Leu Asn Pro Ala Lys Cys Ile Cys
        345                 350                 355

GAA TGT ACA GAA TCT CCC AAT AAA TGT TTC TTA AAA GGA AAA AGA TTT      1577
Glu Cys Thr Glu Ser Pro Asn Lys Cys Phe Leu Lys Gly Lys Arg Phe
360                 365                 370                 375

CAT CAC CAG ACA TGC AGT TGT TAC AGA CCA CCA TGT ACA GTC CGA ACG      1625
His His Gln Thr Cys Ser Cys Tyr Arg Pro Pro Cys Thr Val Arg Thr
                380                 385                 390

AAA CGC TGT GAT GCT GGA TTT CTG TTA GCT GAA GAA GTG TGC CGC TGT      1673
Lys Arg Cys Asp Ala Gly Phe Leu Leu Ala Glu Glu Val Cys Arg Cys
                395                 400                 405

GTA CGC ACA TCT TGG AAA AGA CCA CTT ATG AAT TAAGCGAAGA AAGCACTACT   1726
Val Arg Thr Ser Trp Lys Arg Pro Leu Met Asn
        410                 415

CGCTATATAG TGTCG                                                    1741
```

(2) INFORMATION FOR SEQ ID NO:53:

      (i) SEQUENCE CHARACTERISTICS:
          (A) LENGTH: 418 amino acids
          (B) TYPE: amino acid
          (D) TOPOLOGY: linear

      (ii) MOLECULE TYPE: protein

      (xi) SEQUENCE DESCRIPTION: SEQ ID NO:53:

```
Met His Leu Leu Glu Met Leu Ser Leu Gly Cys Cys Leu Ala Ala Gly
 1               5                  10                  15

Ala Val Leu Leu Gly Pro Arg Gln Pro Pro Val Ala Ala Ala Tyr Glu
            20                  25                  30

Ser Gly His Gly Tyr Tyr Glu Glu Glu Pro Gly Ala Gly Glu Pro Lys
            35                  40                  45
```

```
Ala His Ala Ser Lys Asp Leu Glu Glu Gln Leu Arg Ser Val Ser Ser
     50              55              60

Val Asp Glu Leu Met Thr Val Leu Tyr Pro Glu Tyr Trp Lys Met Phe
 65              70              75                  80

Lys Cys Gln Leu Arg Lys Gly Gly Trp Gln His Asn Arg Glu His Ser
             85              90                  95

Ser Ser Asp Thr Arg Ser Asp Asp Ser Leu Lys Phe Ala Ala Ala His
             100             105             110

Tyr Asn Ala Glu Ile Leu Lys Ser Ile Asp Thr Glu Trp Arg Lys Thr
     115             120             125

Gln Gly Met Pro Arg Glu Val Cys Val Asp Leu Gly Lys Glu Phe Gly
     130             135             140

Ala Thr Thr Asn Thr Phe Phe Lys Pro Pro Cys Val Ser Ile Tyr Arg
145             150             155             160

Cys Gly Gly Cys Cys Asn Ser Glu Gly Leu Gln Cys Met Asn Ile Ser
             165             170             175

Thr Asn Tyr Ile Ser Lys Thr Leu Phe Glu Ile Thr Val Pro Leu Ser
             180             185             190

His Gly Pro Lys Pro Val Thr Val Ser Phe Ala Asn His Thr Ser Cys
     195             200             205

Arg Cys Met Ser Lys Leu Asp Val Tyr Arg Gln Val His Ser Ile Ile
     210             215             220

Arg Arg Ser Leu Pro Ala Thr Gln Thr Gln Cys His Val Ala Asn Lys
225             230             235             240

Thr Cys Pro Lys Asn His Val Trp Asn Asn Gln Ile Cys Arg Cys Leu
             245             250             255

Ala Gln His Asp Phe Gly Phe Ser Ser His Leu Gly Asp Ser Asp Thr
             260             265             270

Ser Glu Gly Phe His Ile Cys Gly Pro Asn Lys Glu Leu Asp Glu Glu
     275             280             285

Thr Cys Gln Cys Val Cys Lys Gly Gly Val Arg Pro Ile Ser Cys Gly
     290             295             300

Pro His Lys Glu Leu Asp Arg Ala Ser Cys Gln Cys Met Cys Lys Asn
305             310             315             320

Lys Leu Leu Pro Ser Ser Cys Gly Pro Asn Lys Glu Phe Asp Glu Glu
             325             330             335

Lys Cys Gln Cys Val Cys Lys Lys Thr Cys Pro Lys His His Pro Leu
             340             345             350

Asn Pro Ala Lys Cys Ile Cys Glu Cys Thr Glu Ser Pro Asn Lys Cys
     355             360             365

Phe Leu Lys Gly Lys Arg Phe His His Gln Thr Cys Ser Cys Tyr Arg
     370             375             380

Pro Pro Cys Thr Val Arg Thr Lys Arg Cys Asp Ala Gly Phe Leu Leu
385             390             395             400

Ala Glu Glu Val Cys Arg Cys Val Arg Thr Ser Trp Lys Arg Pro Leu
             405             410             415

Met Asn
```

(2) INFORMATION FOR SEQ ID NO:54:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 1582 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA (genomic)

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:54:

```
CTAGAGTTGA ACCAGATAAG AAAGTCTCTT CTTCCGGTAA GATATTATGG ACCTATAACA        60

TCTGTGTACT TAAAAGTAGA TTGGGAGTGA AAGGCAGACT TTTGATGTTC TGTACACTGT       120

TGAAACCCCT TAGCGTGGTC CTCTGTAACC TGCTCACCCT GCCCCAAGGA GGCAGCTAGC       180

CAATGCCACC AGCCCAACGG AAACCCCAGT GCTTTTCCAA TGGGGAAATG CAGTCACTTT       240

TCTTTGGATG CTACACATCC TTTCTGGAAT ATGTCTCACA CACATCTCTC TTTATCACCC       300

CCTTTTTCAA GTAAACCAAC TTCTTGCAGA AGCTGACAAT GTGTCTCTTT ACTCTCCACG       360

AAGATTCTGG CCCTTCTCTT CACCTGTCAG AAGTTTAGGA TTCCAAAGGG ATCATTAGCA       420

TCCATCCCAA CAGCCTGCAC TGCATCCTGA GAACTGCGGT TCTTGGATCA TCAGGCAACT       480

TTCAACTACA CAGACCAAGG GAGAGAGGGG ACCCCTCCGA GGTCCCATAG GGTTCTCTGA       540

CATAGTGATG ACCTTTTTCC AAACTTTGAG CAGGGCGCTG GGGGCCAGGC GTGCGGGAGG       600

GAGGACAAGA ACTCGGGAGT GGCCGAGGAT AAAGCGGGGG CTCCCTCCAC CCCACGGTGC       660

CCAGTTTCTC CCCGCTGCAC GTGGTCCAGG GTGGTCGCAT CACCTCTAAA GCCGGTCCCG       720

CCAACCGCCA GCCCCGGGAC TGAACTTGCC CCTCCGGCCG CCCGCTCCCC GCAGGGACA       780

GGGGCGGGGA GGGAGAGATC CAGAGGGGGG CTGGGGGAGG TGGGNCCGCC GGGGAGGAGN       840

CGAGGGAAAC GGGGAGCTCC AGGGAGACGG CTTCCGAGGG AGAGTGAGAG GGGAGGGCAG       900

CCCGGGCTCG GCACGCTCCC TCCCTCGGCC GCTTTCTCTC ACATAAGCGC AGGCAGAGGG       960

CGCGTCAGTC ATGCCCTGCC CCTGCGCCCG CCGCCGCCGC CGCCGCCGCT CAGCCCGGCG      1020

CGCTCTGGAG GATCCTGCGC CGCGGCGCTC CCGGGCCCCG CCGCCGCCAG CCGCCCCGGC      1080

GGCCCTCCTC CCGCCCCCGG CACCGCCGCC AGCGCCCCCG CCGCAGCGCC CGCGGCCCGG      1140

CTCCTCTCAC TTCGGGGAAG GGGAGGGAGG AGGGGGACGA GGGCTCTGGC GGGTTTGGAG      1200

GGGCTGAACA TCGCGGGGTG TTCTGGTGTC CCCCGCCCCG CCTCTCCAAA AAGCTACACC      1260

GACGCGGACC GCGGCGGCGT CCTCCCTCGC CCTCGCTTCA CCTCGCGGGC TCCGAATGCG      1320

GGGAGCTCGG ATGTCCGGTT TCCTGTGAGG CTTTTACCTG ACACCCGCCG CCTTTCCCCG      1380

GCACTGGCTG GGAGGGCGCC CTGCAAAGTT GGGAACGCGG AGCCCCGGAC CCGCTCCCGC      1440

CGCCTCCGGC TCGCCCAGGG GGGGTCGCCG GGAGGAGCCC GGGGGAGAGG GACCAGGAGG      1500

GGCCCGCGGC CTCGCAGGGG CGCCCGCGCC CCCACCCCTG CCCCCGCCAG CGGACCGGTC      1560

CCCCACCCCC GGTCCTTCCA CC                                             1582
```

(2) INFORMATION FOR SEQ ID NO:55:

    (i) SEQUENCE CHARACTERISTICS:

                (A) LENGTH: 20 base pairs
                (B) TYPE: nucleic acid
                (C) STRANDEDNESS: single
                (D) TOPOLOGY: linear

        (ii) MOLECULE TYPE: DNA

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:55:

CACGGCTTAT GCAAGCAAAG   20

(2) INFORMATION FOR SEQ ID NO:56:

        (i) SEQUENCE CHARACTERISTICS:
                (A) LENGTH: 20 base pairs
                (B) TYPE: nucleic acid
                (C) STRANDEDNESS: single
                (D) TOPOLOGY: linear

        (ii) MOLECULE TYPE: DNA

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:56:

AACACAGTTT TCCATAATAG   20

(2) INFORMATION FOR SEQ ID NO:57:

        (i) SEQUENCE CHARACTERISTICS:
                (A) LENGTH: 18 base pairs
                (B) TYPE: nucleic acid
                (C) STRANDEDNESS: single
                (D) TOPOLOGY: linear

        (ii) MOLECULE TYPE: DNA

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:57:

GCCACGGTAG GTCTGCGT   18

(2) INFORMATION FOR SEQ ID NO:58:

        (i) SEQUENCE CHARACTERISTICS:
                (A) LENGTH: 18 base pairs
                (B) TYPE: nucleic acid
                (C) STRANDEDNESS: single
                (D) TOPOLOGY: linear

        (ii) MOLECULE TYPE: DNA

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:58:

TTTCTTTGAC AGGCTTAT   18

(2) INFORMATION FOR SEQ ID NO:59:

        (i) SEQUENCE CHARACTERISTICS:
                (A) LENGTH: 21 base pairs
                (B) TYPE: nucleic acid
                (C) STRANDEDNESS: single
                (D) TOPOLOGY: linear

        (ii) MOLECULE TYPE: DNA

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:59:

ATCTTGAAAA GTAAGTATGG G   21

(2) INFORMATION FOR SEQ ID NO:60:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 20 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:60:

ATGACTTGAC AGGTATTGAT   20

(2) INFORMATION FOR SEQ ID NO:61:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 20 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:61:

AGCAAGACGG TGGGTATTGT   20

(2) INFORMATION FOR SEQ ID NO:62:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 22 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:62:

CCCTTCTTTG TAGTTATTTG AA   22

(2) INFORMATION FOR SEQ ID NO:63:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 20 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:63:

CCACAGTGAG TATGAATTAA   20

(2) INFORMATION FOR SEQ ID NO:64:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 18 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:64:

TTCTTCCAAA GGTGTCAG   18

    (2) INFORMATION FOR SEQ ID NO:65:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 18 base pairs
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: single
            (D) TOPOLOGY: linear

        (ii) MOLECULE TYPE: DNA

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:65:

GGAGATGGTA GCAGAATG   18

    (2) INFORMATION FOR SEQ ID NO:66:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 23 base pairs
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: single
            (D) TOPOLOGY: linear

        (ii) MOLECULE TYPE: DNA

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:66:

CTATTTGTCT AGACTCAACA GAT   23

    (2) INFORMATION FOR SEQ ID NO:67:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 22 base pairs
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: single
            (D) TOPOLOGY: linear

        (ii) MOLECULE TYPE: DNA

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:67:

CAAACATGCA GGTAAGAGAT CC   22

    (2) INFORMATION FOR SEQ ID NO:68:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 21 base pairs
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: single
            (D) TOPOLOGY: linear

        (ii) MOLECULE TYPE: DNA

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:68:

TGTTCTCCTA GCTGTTACAG A   21

    (2) INFORMATION FOR SEQ ID NO:69:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 24 base pairs
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: single
            (D) TOPOLOGY: linear

        (ii) MOLECULE TYPE: DNA

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:69:

GGCGAGGTCA AGGTAGGTGC AAGG   24

(2) INFORMATION FOR SEQ ID NO:70:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 26 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:70:

ATTGTCTTTG ACAGGCTTTT TGAAGG   26

(2) INFORMATION FOR SEQ ID NO:71:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 21 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:71:

GAGATCCTGA AAAGTAAGTA G   21

(2) INFORMATION FOR SEQ ID NO:72:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 24 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:72:

TGTGACTCGA CAGGTATTGA TAAT   24

(2) INFORMATION FOR SEQ ID NO:73:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 20 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:73:

CTCAGCAAGA CGGTAGGTAT   20

(2) INFORMATION FOR SEQ ID NO:74:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 25 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:74:

TTGTCCCTTG TAGTTGTTTG AAATT   25

(2) INFORMATION FOR SEQ ID NO:75:

        (i) SEQUENCE CHARACTERISTICS:
             (A) LENGTH: 20 base pairs
             (B) TYPE: nucleic acid
             (C) STRANDEDNESS: single
             (D) TOPOLOGY: linear

        (ii) MOLECULE TYPE: DNA

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:75:

ACATTACCAC AGTGAGTATG   20

(2) INFORMATION FOR SEQ ID NO:76:

        (i) SEQUENCE CHARACTERISTICS:
             (A) LENGTH: 26 base pairs
             (B) TYPE: nucleic acid
             (C) STRANDEDNESS: single
             (D) TOPOLOGY: linear

        (ii) MOLECULE TYPE: DNA

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:76:

GTCTCCCCAA AAGGTGTCAG GCAGCT   26

(2) INFORMATION FOR SEQ ID NO:77:

        (i) SEQUENCE CHARACTERISTICS:
             (A) LENGTH: 23 base pairs
             (B) TYPE: nucleic acid
             (C) STRANDEDNESS: single
             (D) TOPOLOGY: linear

        (ii) MOLECULE TYPE: DNA

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:77:

AATGTTGAAG ATGGTAAGTA AAA   23

(2) INFORMATION FOR SEQ ID NO:78:

        (i) SEQUENCE CHARACTERISTICS:
             (A) LENGTH: 16 base pairs
             (B) TYPE: nucleic acid
             (C) STRANDEDNESS: single
             (D) TOPOLOGY: linear

        (ii) MOLECULE TYPE: DNA

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:78:

TCTAGACTCA ACCAAT   16

(2) INFORMATION FOR SEQ ID NO:79:

        (i) SEQUENCE CHARACTERISTICS:
             (A) LENGTH: 22 base pairs
             (B) TYPE: nucleic acid
             (C) STRANDEDNESS: single
             (D) TOPOLOGY: linear

        (ii) MOLECULE TYPE: DNA

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:79:

CAAACATGCA GGTAAGGAGT GT   22

(2) INFORMATION FOR SEQ ID NO:80:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 24 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:80:

TTTTCCCCTA GTTGTTACAG AAGA   24

(2) INFORMATION FOR SEQ ID NO:81:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 19 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:81:

Leu Ser Lys Thr Val Ser Gly Ser Glu Gln Asp Leu Pro His Glu Leu
1   5   10   15

His Val Glu

- 135 -

CLAIMS

1. A purified and isolated polypeptide which is capable of binding to Flt4 receptor tyrosine kinase.

2. A mammalian polypeptide according to claim 1.

3. A human polypeptide according to claim 1.

4. A polypeptide according to claim 3 having an apparent molecular weight of about 32 kD as assessed by SDS-PAGE under reducing conditions.

5. A polypeptide according to claim 3 that is substantially free of other human polypeptides.

6. A purified and isolated polypeptide according to claim 1, said polypeptide being encoded by plasmid pFLT4-L, deposited as ATCC Accession Number 97231.

7. A polypeptide according to claim 1 having an amino acid sequence comprising a portion of SEQ ID NO: 33.

8. A polypeptide according to claim 1 comprising an amino acid sequence set forth in SEQ ID NO: 33 from about residue 161 of SEQ ID NO: 33 to about residue 211 of SEQ ID NO: 33.

9. A polypeptide according to claim 1 comprising an amino acid sequence set forth in SEQ ID NO: 33 from about residue 131 of SEQ ID NO: 33 to about residue 211 of SEQ ID NO: 33.

10. A polypeptide according to claim 1 comprising an amino acid sequence set forth in SEQ ID NO: 33 from residue 113 of SEQ ID NO: 33 to residue 213 of SEQ ID NO: 33.

- 136 -

11. A polypeptide according to claim 1 comprising an amino acid sequence set forth in SEQ ID NO: 33 from residue 113 of SEQ ID NO: 33 to residue 227 of SEQ ID NO: 33.

12. A polypeptide according to claim 1 comprising amino acids 103 to 217 of SEQ ID NO: 33.

13. A polypeptide according to claim 1 comprising amino acids 103 to 225 of SEQ ID NO: 33.

14. A polypeptide according to claim 1 comprising amino acids 103 to 227 of SEQ ID NO: 33.

15. A polypeptide according to claim 1 comprising amino acids 32 to 227 of SEQ ID NO: 33.

16. A polypeptide according to any of claims 1-15, wherein said polypeptide is capable of stimulating tyrosine phosphorylation of Flt4 receptor tyrosine kinase in a host cell expressing said Flt4 receptor tyrosine kinase.

17. A polypeptide according to claim 1 having the amino acid sequence of residues 1 to 419 of SEQ ID NO: 33.

18. A polypeptide according to any of claims 1-17 further comprising a detectable label.

19. A purified protein comprising a first polypeptide linked to a second polypeptide, wherein at least one of said first polypeptide and said second polypeptide is a polypeptide according to any of claims 1-17, and wherein said protein is capable of binding to Flt4 receptor tyrosine kinase.

20. A purified protein according to claim 19 wherein said first polypeptide is covalently linked to said

second polypeptide.

21.   A purified protein according to claim 19 whe-
rein each of said first polypeptide and said second poly-
peptide is a polypeptide according to any of claims 1-17.

22.   A purified and isolated nucleic acid comp-
rising a nucleotide sequence encoding a polypeptide accor-
ding to any of claims 1-17.

23.   A nucleic acid according to claim 22 comp-
rising a nucleotide sequence encoding a polypeptide consis-
ting of amino acids 113 to 213 of SEQ ID NO: 33.

24.   A nucleic acid according to claim 22 comp-
rising a nucleotide sequence encoding a polypeptide consis-
ting of amino acids 103 to 217 of SEQ ID NO: 33.

25.   A nucleic acid according to claim 22 having
the nucleotide sequence of nucleotides 352 to 1608 of SEQ
ID NO: 32.

26.   A nucleic acid according to claim 22 comp-
rising a VEGF-C encoding insert of plasmid pFLT4-L, de-
posited as ATCC Accession Number 97321.

27.   A vector comprising a nucleic acid according
to any of claims 22-26.

28.   A host cell transformed or transfected with a
nucleic acid according to any of claims 22-26.

29.   An antibody which is specifically reactive
with a polypeptide according to any of claims 1-17.

- 138 -

30. An antibody according to claim 29 which is a monoclonal antibody.

31. A pharmaceutical composition comprising a polypeptide according to any of claims 1-17 in a pharmaceutically acceptable diluent, adjuvant, excipient, or carrier.

32. A method of making a polypeptide capable of specifically binding to Flt4 receptor tyrosine kinase, said method comprising the steps of:
(a) expressing a nucleic acid according to any of claims 22-26 in a host cell; and
(b) purifying a polypeptide capable of specifically binding to Flt4 receptor tyrosine kinase from said host cell or from a growth medium of said host cell.

33. A polypeptide capable of specifically binding to Flt4 receptor tyrosine kinase, said polypeptide produced by the method according to claim 32.

34. A murine polypeptide according to claim 1.

35. A polypeptide according to claim 34, said polypeptide comprising a portion of the amino acid sequence set forth in SEQ ID NO: 41, said portion being capable of specifically binding to an Flt4 receptor tyrosine kinase.

36. A purified and isolated nucleic acid encoding the polypeptide according to claim 34 or 35.

37. A purified and isolated nucleic acid having at least about 16 nucleotides, said nucleic acid specifically hybridizing to a human gene encoding VEGF-C.

- 139 -

38.   A nucleic acid according to claim 37 which
hybridizes to a human gene encoding VEGF-C, under hybridi-
zation conditions wherein said nucleic acid fails to hybri-
dize to a human gene encoding VEGF or VEGF-B, and wherein
5   said nucleic acid comprises a continuous nucleotide sequen-
ce of at least twenty nucleotides from a nucleotide sequen-
ce selected from the group consisting of:
          SEQ ID NO: 32, and
          a nucleotide sequence complementary to SEQ ID NO:
10   32.

39.   A method for detecting endothelial cells in a
biological tissue comprising the steps of:
          exposing a biological tissue comprising endothe-
15   lial cells to a polypeptide according to any of claims 1-
18, under conditions wherein said polypeptide binds to en-
dothelial cells; and
          detecting said polypeptide bound to endothelial
cells in said biological tissue, thereby detecting said
20   endothelial cells.

40.   The method according to claim 39, further
comprising the step of washing said biological tissue, said
washing step being performed after said exposing step and
25   before said detecting step.

41.   A method of modulating the growth of mammali-
an endothelial cells comprising the steps of:
          exposing mammalian endothelial cells to a polypep-
30   tide according to any of claims 1-17 in an amount effective
to modulate the growth of mammalian endothelial cells.

42.   A purified and isolated polypeptide which is
capable of binding to Kdr (VEGFR-2), said polypeptide ha-
35   ving an amino acid sequence comprising a portion of SEQ ID
NO: 33.

- 140 -

43. A purified nucleic acid comprising a promoter
for VEGF-C.

44. A nucleic acid according to claim 43 comp-
rising a portion of SEQ ID NO: 54, wherein said portion is
capable of promoting expression of a protein encoding gene
operatively linked thereto under conditions wherein VEGF-C
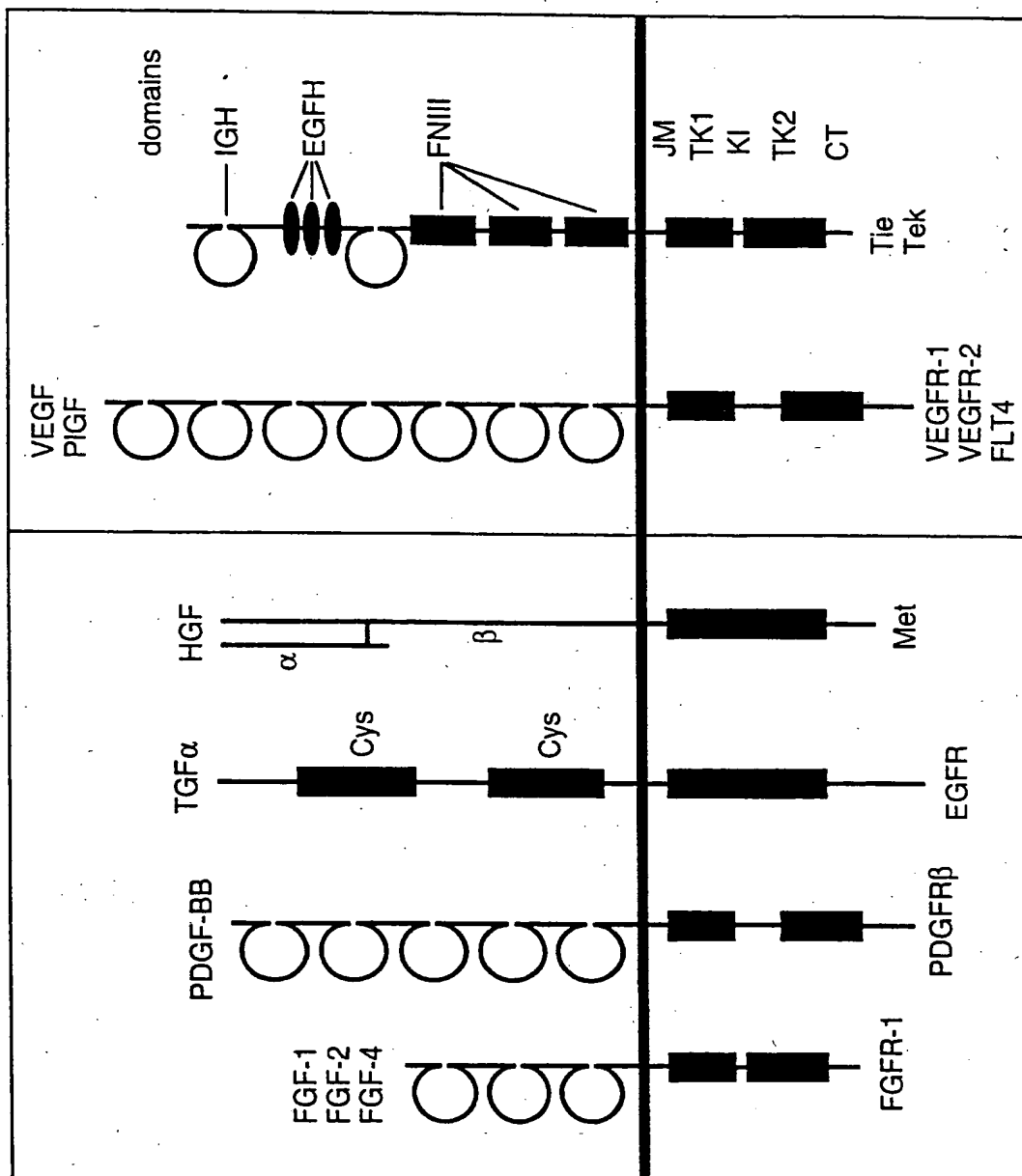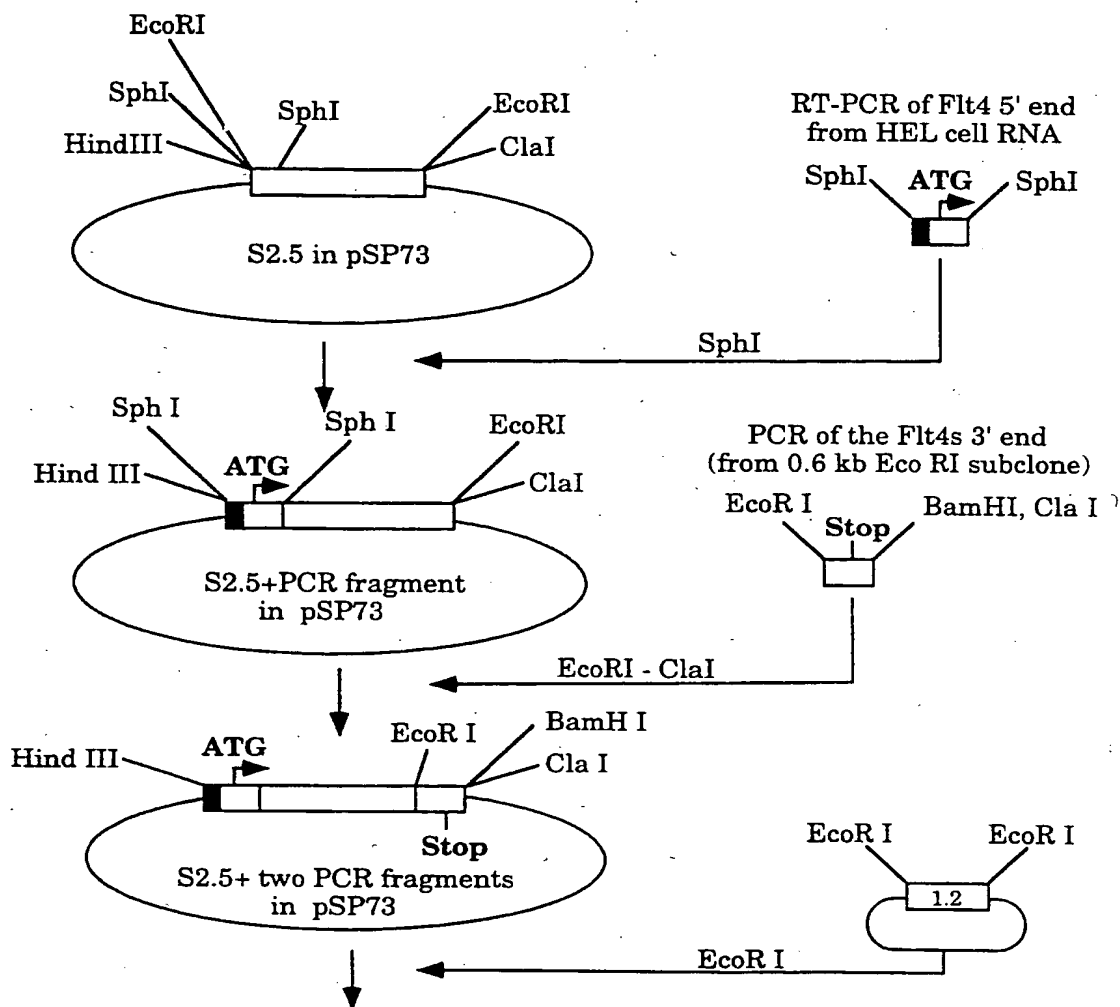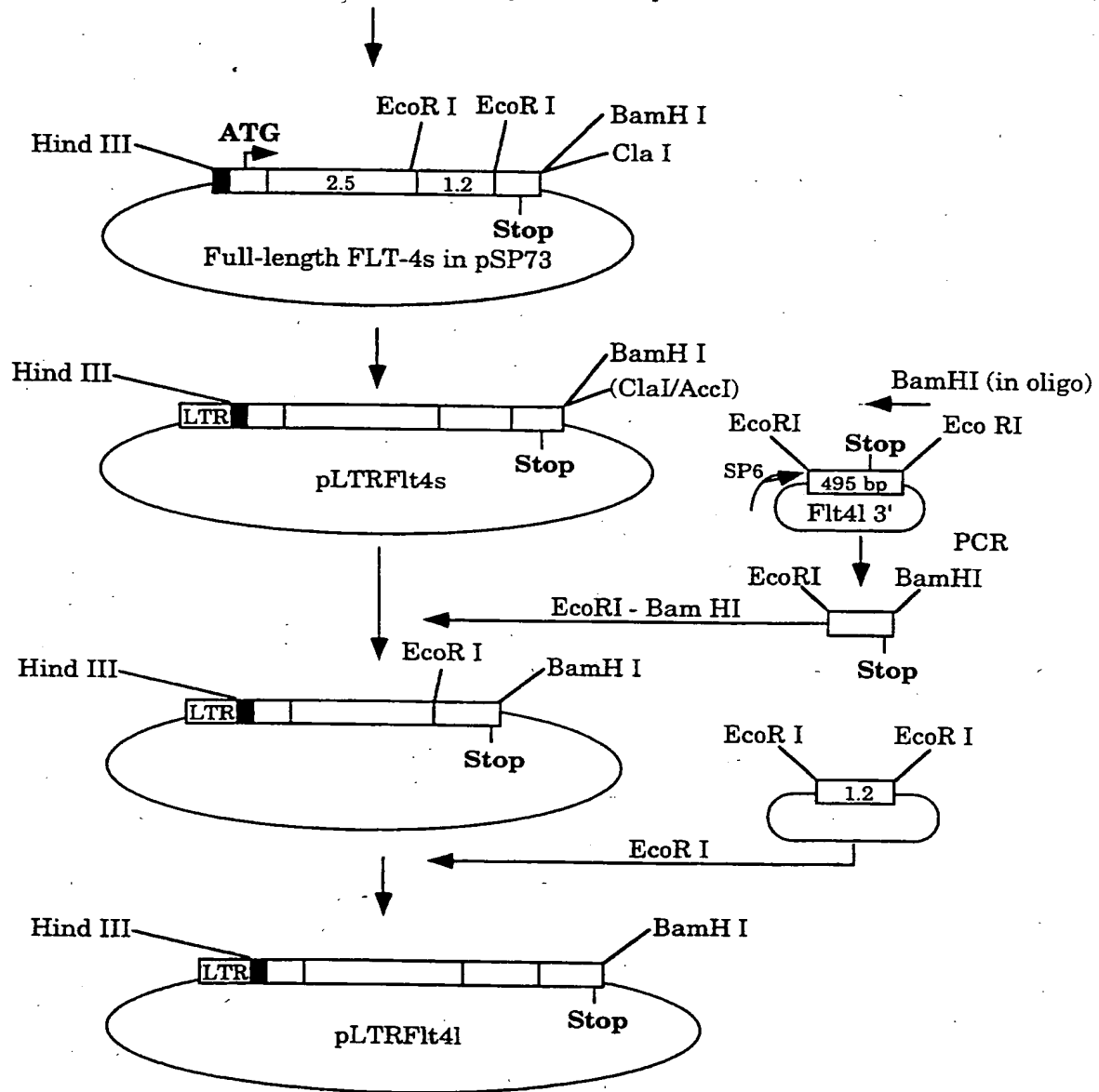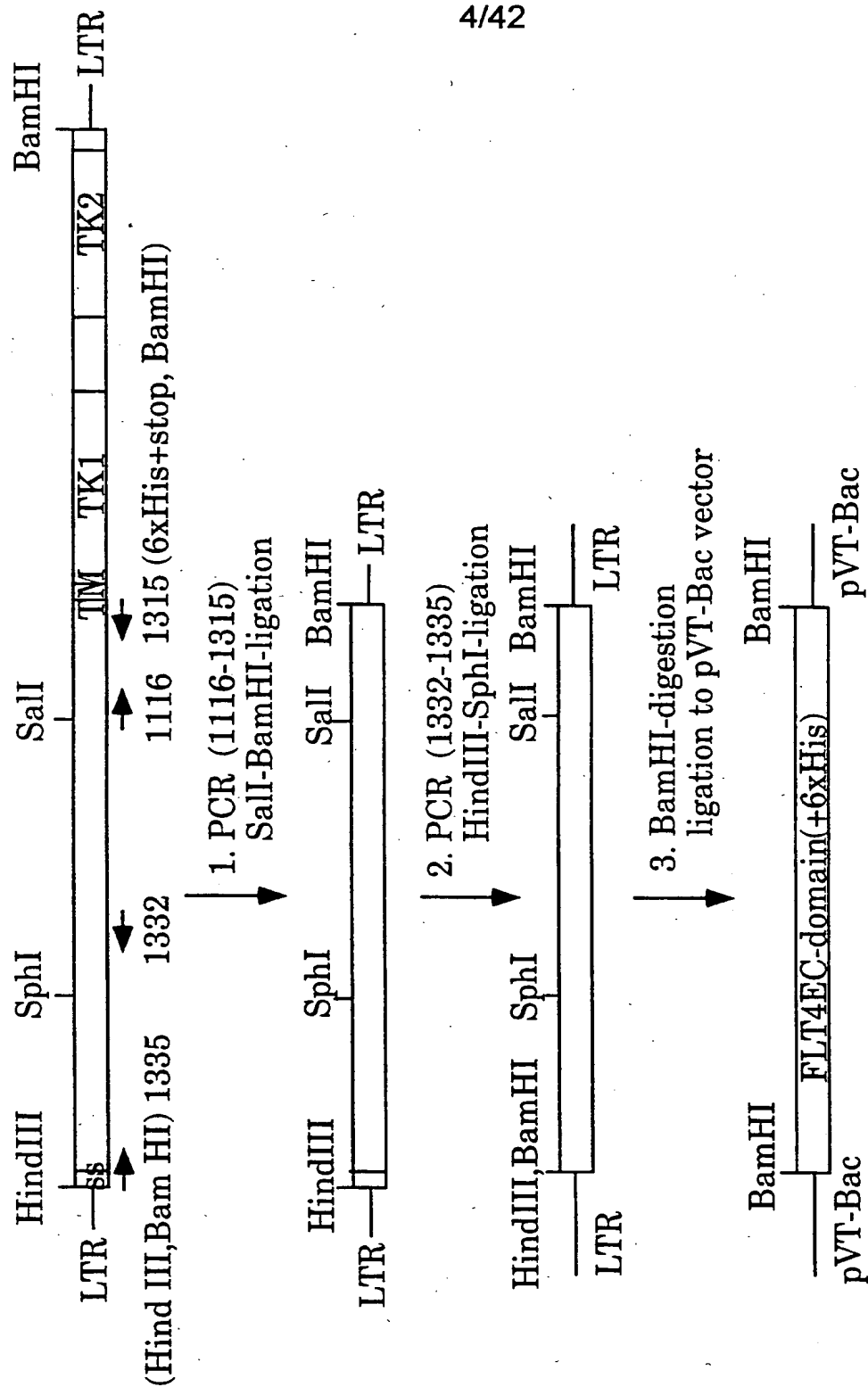is expressed in native host cells.

Figure 1

Figure 2 (1 of 2)

Figure 2 (2 of 2)

Figure 3

Figure 4

A   **B**

anti-PTyr   anti-Flt4   anti-PTyr

   1   2   3   m.w.   4   5   6   m.w.   1   2

─ 170 ─   ─ 170 ─

─ 116 ─   ─ 116 ─

PerVO4 PC-3 CM M   PerVO4 PC-3 CM M   PerVO4 PC-3 CM

C   anti-PTyr

         1   2   3   4   5   m.w.   6   7   8   9

                              ─ 170 ─

                              ─ 116 ─

PC-3 CM   -   +   +   -   +   -   +   -   +

Heparin-   -   -   +   -   -
Sepharose   ─────────────   ──────   NIH 3T3/   NIH 3T3/
                NIH 3T3/   NIH 3T3   Flt4   FGFR-4
                Flt4
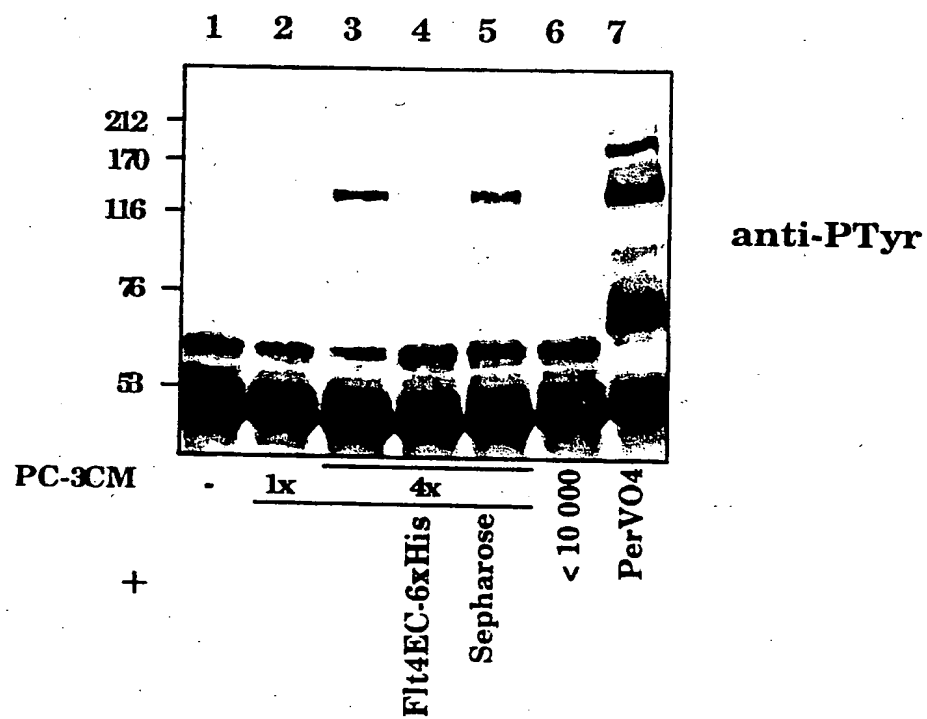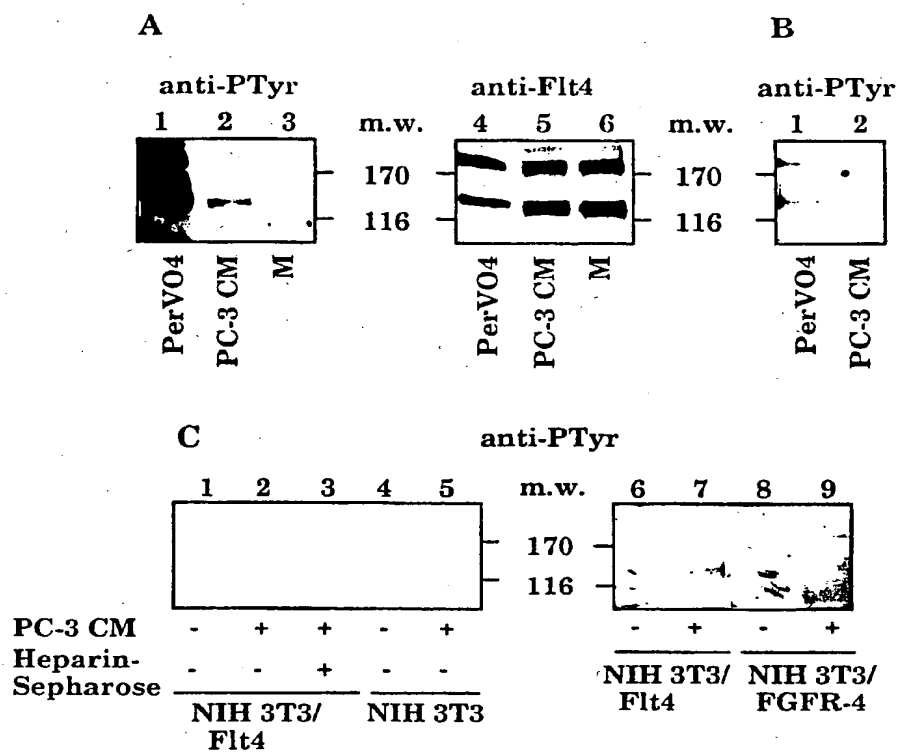
Figure 5

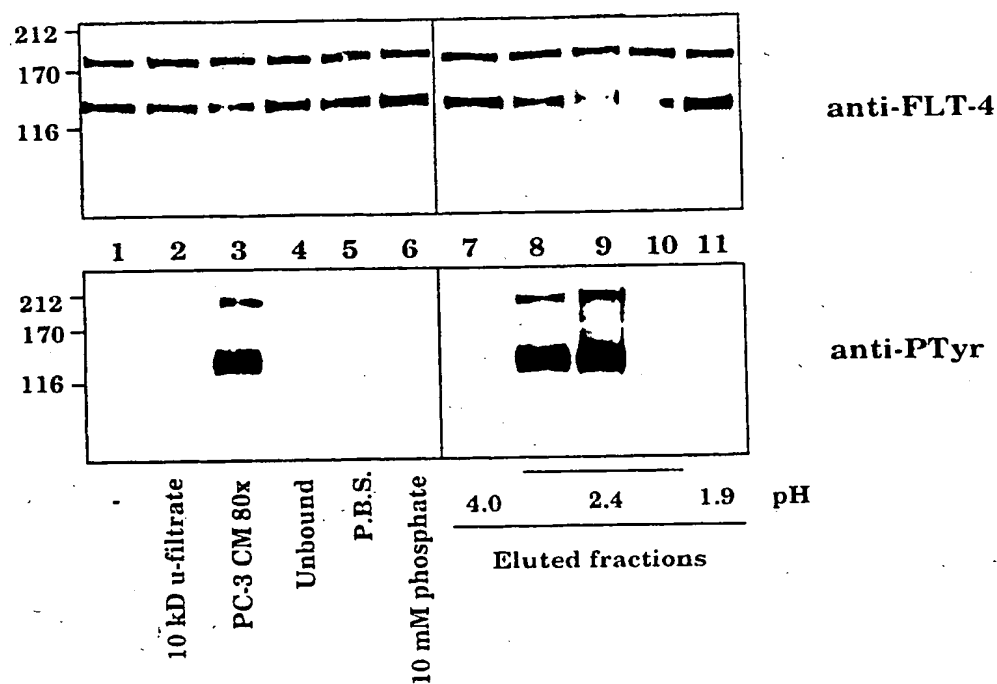Figure 6

Figure 7

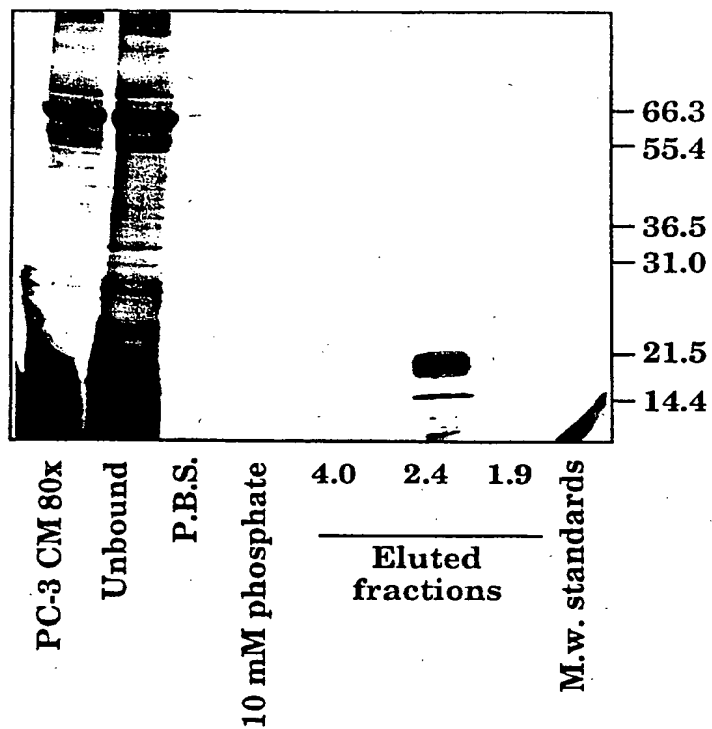**VEGF-C alignment**

```
      1                                                                 50
Hum   HMLLGFFSVA CSLLAAALLP GPREAPAAAA AFESGLDLSD AEPDAGEATA
Mou   MHLLCFLSLA CSLLAAALIP SPREAPATVA AFESGLGFSE AEPDGGEVKA
Qua   MHLLEMLSLG CCLAAGAVLL GPRQPPVA.A AYESGHGYYE EEPGAGEPKA


      51                                                               100
Hum   YASKDLEEQL RSVSSVDELM TVLYPEYWKM YKCQLRKGGW QHNREQANLN
Mou   FEGKDLEEQL RSVSSVDELM SVLYPDYWKM YKCQLRKGGW Q....QPTLN
Qua   HASKDLEEQL RSVSSVDELM TVLYPEYWKM FKCQLRKGGW QHNREHSSSD


      101                                                              150
Hum   SRTEETIKFA AAHYNTEILK SIDNEWRKTQ CMPREVCIDV GKEFGVATNT
Mou   TRTGDSVKFA AAHYNTEILK SIDNEWRKTQ CMPREVCIDV GKEFGAATNT
Qua   TRSDDSLKFA AAHYNAEILK SIDTEWRKTQ GMPREVCVDL GKEFGATTNT


      151                                                              200
Hum   FFKPPCVSVY RCGGCCNSEG LQCMNTSTSY LSKTLFEITV PLSQGPKPVT
Mou   FFKPPCVSVY RCGGCCNREG LQCMNTSTGY LSKTLFEITV PLSQGPKPVT
Qua   FFKPPCVSIY RCGGCCNSEG LQCMNISTNY ISKTLFEITV PLSHGPKPVT


      201                                                              250
Hum   ISFANHTSCR CMSKLDVYRQ VHSIIRRSLP ATLPQCQAAN KTCPTNYMWN
Mou   ISFANHTSCR CMSKLDVYRQ VHSIIRRSLP ATLPQCQAAN KTCPTNYVWN
Qua   VSFANHTSCR CMSKLDVYRQ VHSIIRRSLP ATQTQCHVAN KTCPKNHVWN


      251                                                              300
Hum   NHICRCLAQE DFMFSSDAGD DSTDGFHDIC GPNKELDEET CQCVCRAGLR
Mou   NYMCRCLAQQ DFIFYSNVED DSTNGFHDVC GPNKELDEDT CQCVCKGGLR
Qua   NQICRCLAQH DFGFSSHLGD SDTSEGFHIC GPNKELDEET CQCVCKGGVR


      301                                                              350
Hum   PASCGPHKEL DRNSCQCVCK NKLFPSQCGA NREFDENTCQ CVCKRTCPRN
Mou   PSSCGPHKEL DRDSCQCVCK NKLFPNSCGA NREFDENTCQ CVCKRTCPRN
Qua   PISCGPHKEL DRASCQCMCK NKLLPSSCGP NKEFDEEKCQ CVCKKTCPKH


      351                                                              400
Hum   QPLNPGKCAC ECTESPQKCL LKGKKFHHQT CSCYRRPCTN RQKACEPGFS
Mou   QPLNPGKCAC ECTENTQKCF LKGKKFHHQT CSCYRRPCAN RLKHCDPGLS
Qua   HPLNPAKCIC ECTESPNKCF LKGKRFHHQT CSCYRPPCTV RTKRCDAGFL


      401            420
Hum   YSEEVCRCVP SYWKRPQMS*
Mou   FSEEVCRCVP SYWKRPHLN.
Qua   LAEEVCRCVR TSWKRPLMN*
```
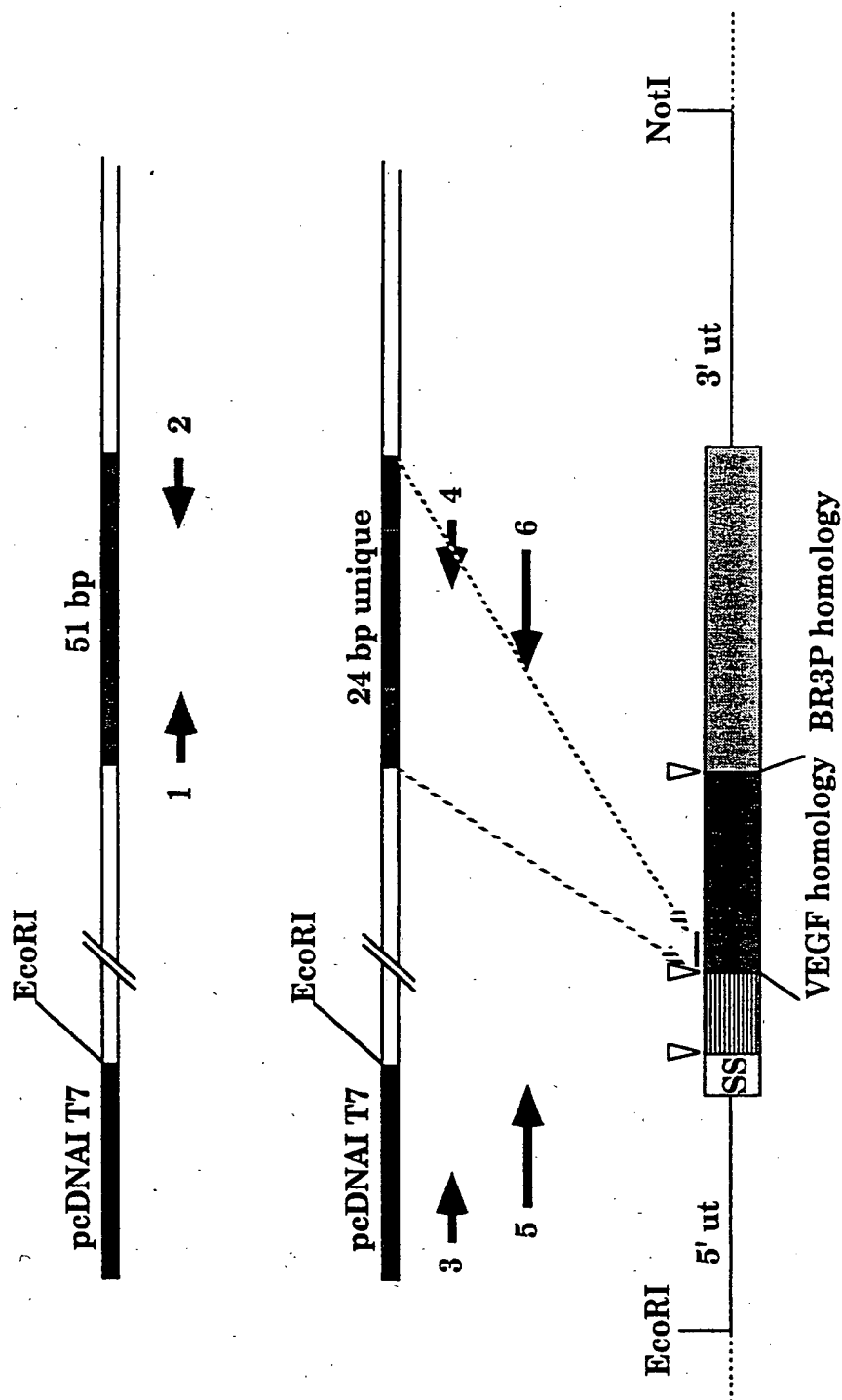
FIGURE 8

Figure 9

```
                                                               50
PDGF-A     .......... .......... .......... .......... .MRTLACLLL
PDGF-B     .......... .......... .......... .......... MNRCWA.LFL
P1GF-1     .......... .......... .......... .......... ..........
VEGF165    .......... .......... .......... .......... ..........
VEGF-B167  .......... .......... .......... .......... ..........
VEGF-C     MHLLGFFSVA CSLLAAALLP GPREAPAAAA AFESGLDLSD AEPDAGEATA

           51                                                 100
PDGF-A     LGCGYLAHVL AEEAEIPREV IERLARSQIH SIRDLQRLLE IDSVGSEDSL
PDGF-B     SLCCYLRLVS AEGDPIPEEL YEMLSDHSIR SFDDLQRLLH GDP.GEEDGA
P1GF-1     .......... .....MPVM  RLFPC..FLQ LLAGLAL... PAVPPQQW..
VEGF165    .......... .........M NFLLS..WVH WSLALLLYLH HAKWSQAA..
VEGF-B167  .......... ......M    SPLLR..RLL LAALLQLAPA QAPVSQP...
VEGF-C     YASKDLEEQL RSVSSVDELM TVLYPEYWKM YKCQLRKGGW QHNREQANLN

           101                                                150
PDGF-A     DTSLRAHGVH ATKHVPEKRP LPIRRKRSI. .....EEAVP AVCKTRTVIY
PDGF-B     ELDLNMTRSH SGGELES... .LARGRRSLG SLTIAEPAMI AECKTRTEVF
P1GF-1     .....ALSAG NGSSEVEVVP FQE.VWGR.. .......... SYCRALERLV
VEGF165    .....PMAEG GGQNHHEVVK FMD.VYQR.. .......... SYCHPIETLV
VEGF-B167  .......D   APGHQRKVVS WID.VYTR.. .......... ATCQPREVVV
VEGF-C     SRTEETIKFA AAHYNTEILK SIDNEWRK.. .......... TQCMPREVCI

           151                                                200
PDGF-A     EIPRSQVDPT SANFLIWPPC VEVKRCTGCC NTSSVKCQPS RVHHRSVKVA
PDGF-B     EISRRLIDRT NANFLVWPPC VEVQRCSGCC NNRNVQCRPT QVQLRPVQVR
P1GF-1     DVVSEYPSEV ..EHMFSPSC VSLLRCTGCC GDENLHCVPV ETANVTMQLL
VEGF165    DIFQEYPDEI ..EYIFKPSC VPLMRCGGCC NDEGLECVPT EESNITMQIM
VEGF-B167  PLTVELMGTV ..AKQLVPSC VTVQRCGGCC PDDGLECVPT GQHQVRMQIL
VEGF-C     DVGKEFGVAT ..NTFFKPPC VSVYRCGGCC NSEGLQCMNT STSYLSKTLF
```

FIG. 10 (1 of 3)

```
           201                                             250
PDGF-A     KVEYVRKKPK LKEVQVRLEE HLECACAT.. .......... TSLNPDYREE
PDGF-B     KIEIVRKKPI FKKATVTLED HLACKCETVA AARPVTRSPG GSQEQRAKTP
PlGF-1     KIRSG..DRP .SYVELTFSQ HVRCECRPLR EK.........
VEGF165    RIKPH..QGQ .HIGEMSFLQ HNKCECRPKK DR.........
VEGF-B167  MIRYP..SSQ ..LGEMSLEE HSQCECRPKK KD.........
VEGF-C     EITVPLSQGP .KPVTISFAN HTSCRCMSKL DVYRQVHSII RRSLPATLPQ

           251                                             300
PDGF-A     DTDVR..... .......... .......... .......... ..........
PDGF-B     QTRVTIRTVR VRRPPKGKHR KFKHTHDKTA LKETLGA...
PlGF-1     .......... .......... .......... .MKPERCGDA VPRR......
VEGF165    .......... .......... .......... .ARQENPCGP CSERRKHLFV
VEGF-B167  .......... .......... .........S AVKPDSPRPL CPRCTQHHQR
VEGF-C     CQAANKTCPT NYMWNNHICR CLAQEDFMFS SDAGDDSTDG FHDICGPNKE

           301                                             350
PDGF-A     .......... .......... .......... .......... ..........
PDGF-B     .......... .......... .......... .......... ..........
PlGF-1     .......... .......... .......... .......... ..........
VEGF165    QDPQTCKCSC KNTDS.RCKA RQLELNERTC RCDKPRR...
VEGF-B167  PDPRTCRCRC RRRSFLRCQG RGLELNPDTC RCRKLRR...
VEGF-C     LDEETCQCVC RAGLRPASCG PHKELDRNSC QCVCKNKLFP SQCGANREFD
```

FIG. 10 (2 of 3)

```
                 351                                                400
        PDGF-A   .......... .......... .......... .......... ..........
        PDGF-B   .......... .......... .......... .......... ..........
        P1GF-1   .......... .......... .......... .......... ..........
        VEGF165  .......... .......... .......... .......... ..........
        VEGF-B167 ......... .......... .......... .......... ..........
        VEGF-C   ENTCQCVCKR TCPRNQPLNP GKCACECTES PQKCLLKGKK FHHQTCSCYR

                 401                                      434
        PDGF-A   .......... .......... .......... ....
        PDGF-B   .......... .......... .......... ....
        P1GF-1   .......... .......... .......... ....
        VEGF165  .......... .......... .......... ....
        VEGF-B167 ......... .......... .......... ....
        VEGF-C   RPCTNRQKAC EPGFSYSEEV CRCVPSYWKR PQMS
```

FIG. 10 (3 of 3)

14/42

FIGURE 11

Figure 12

A

Figure 13A

B
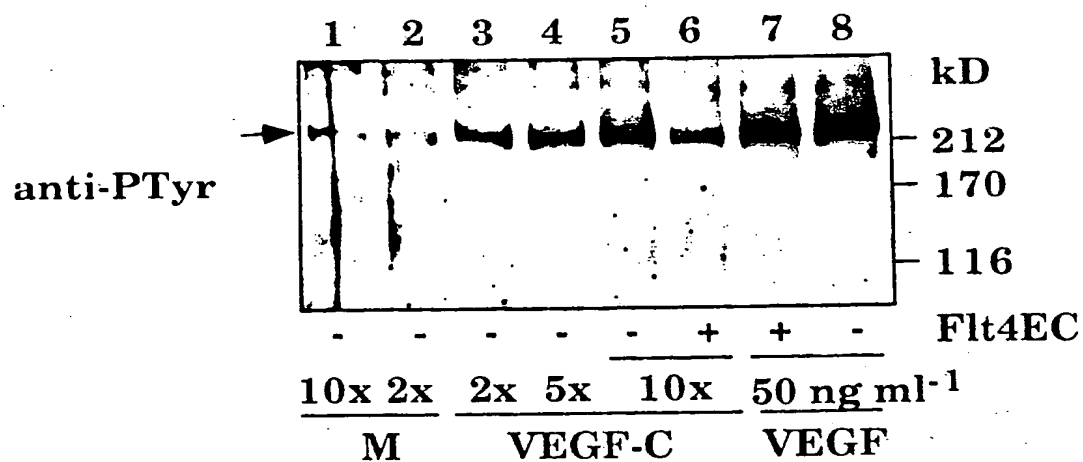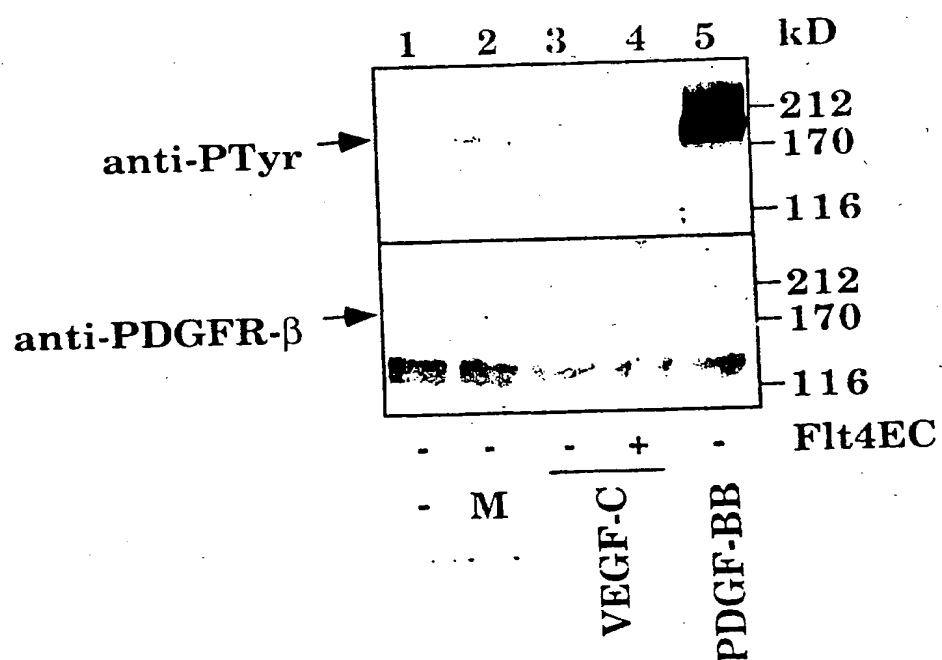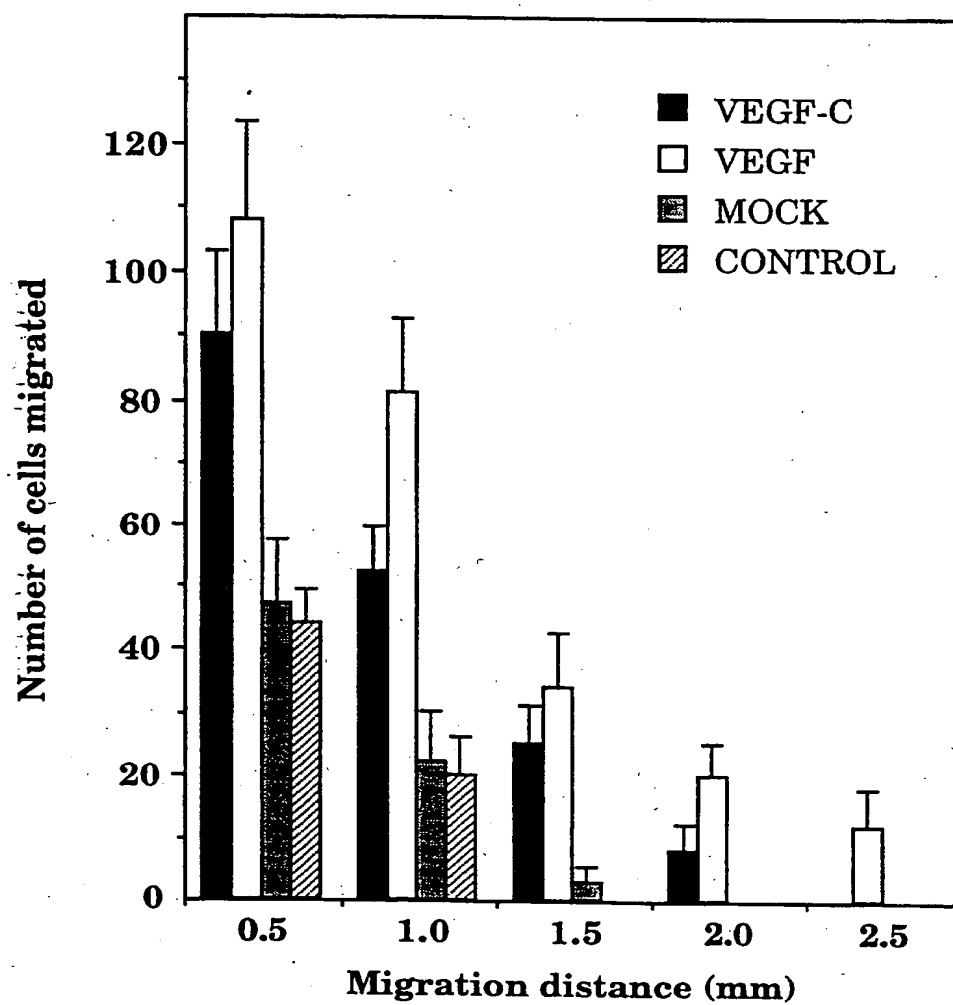
Figure 13B

Figure 14A
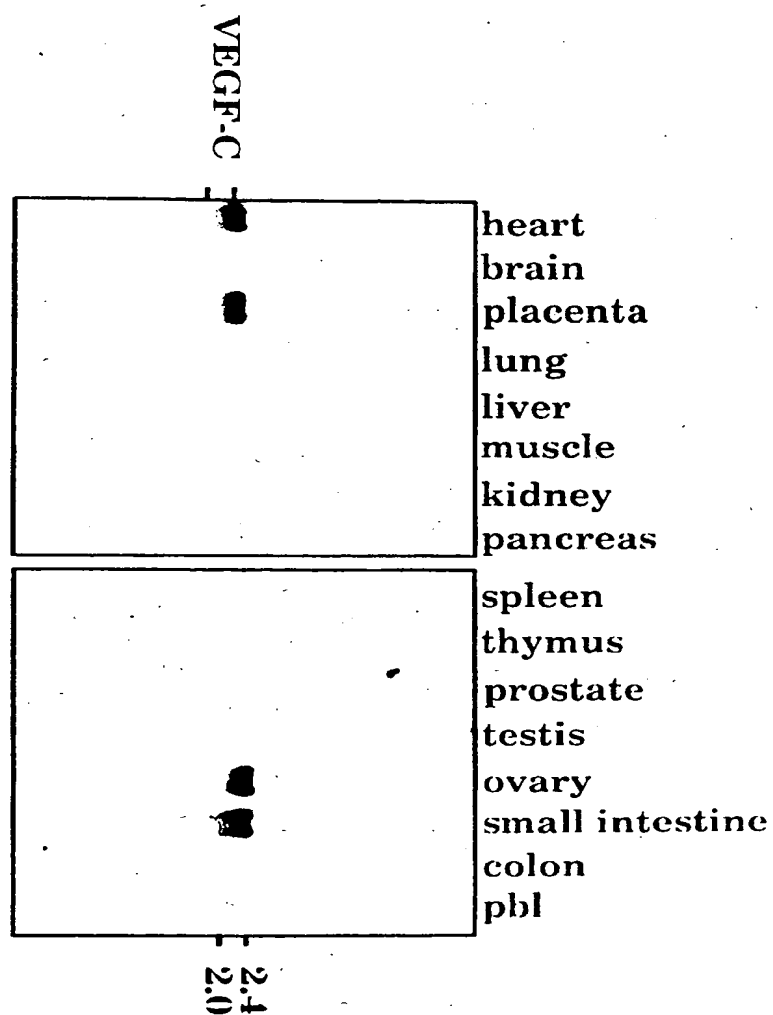
Figure 14B

Figure 15

Figure 16A
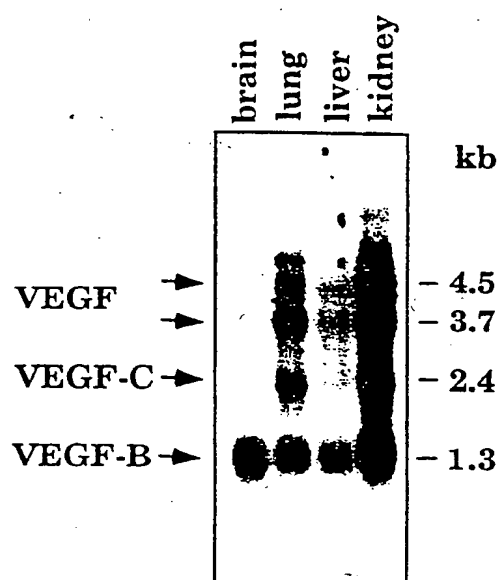
FIG. 16B

HUMAN

Exon length	Donor site	Intron length	Acceptor site

```
                    G...E...A...T(49)..................................A...Y...A...S.
E1..........GGC.GAG.GCC.ACG.gtaggtctgcgt...>10.kb..TTTCTTTGACAG.GCT.TAT.GCA.AGC
            E...I...L...K(116)....................................S...I...D...N.
E2.214.bp..GAG.ATC.TTG.AAA.Agtaagtatggg.....4.kb....atgacttgacagGT.ATT.GAT.AAT
            L...S...K...T(180).....................................L...F...E...I.
E3.191.bp..CTC.AGC.AAG.ACG.gtgggtattgt......9.kb.cccttcttgtag.TTA.TTT.GAA.ATT
            T...L...P...Q(231)..................................C...Q...A...A.
E4.152.bp..ACA.CTA.CCA.CAgtgagtatgaattaaa.>10.kb..ttcttccaaagG.TGT.CAG.GCA.GCG
            A...G...D...(266)....................................D...S...T...D.
E5.107.bp..GCT.GGA.GAT.Ggtagcagaatg.......301.bp....ctatttgtctagAC.TCA.ACA.GAT
            Q...T...C...S(378)..................................C...Y...R...R.
E6.334.bp..CAA.ACA.TGC.AGgtaagagatcc......>10.kb..tgttctcctagC.TGT.TAC.AGA.CGG
            Q...M...S(415)Stop...............
E7.(501).bp..CAA.ATG.AGC.TAA.GTATGTACTGTT..ATTGTATTAT
```

Figure 17 (1 of 2)

MOUSE

Exon length	Donor site	Intron length	Acceptor site
	......G...E...V...K(49)............	....	A...F...E...G.
E1.........	GGC.GAG.GTC.AAG.gtaggtgcaagg	>10.kb.attgtctttgacag.	GCT.TTT.TGA.AGG
	......E...I...L...K(116).......	....	S...I...D...N.
E2.201.bp.	GAG.ATC.CTG.AAA.Agtaagtag......	.4.kb....tgtgactcgacag	GT.ATT.GAT.AAT
	...L...S...K...T(180).........	....	L...F...E...I.
E3.191.bp.	CTC.AGC.AAG.ACG.gtaggtat.......	.9.kb..ttgtcccttgtag.	TTG.TTT.GAA.ATT
	...T...L...P...Q(231).........	....	C...Q...A...A.
E4.152.bp.	ACA.TTA.CCA.CAgtgagtatg........	10.kb.gtctcccaaaag	G.TGT.CAG.GCA.GCT
	...N...V...E...D(266).........	....	D...S...T...N.
E5.107.bp.	AAT.GTT.GAA.GAT.Ggtaagtaaaa...	350.bp...........tctag	AC.TCA.ACC.AAT
	...Q...T...C...S(378).........	....	C...Y...R...R.
E6.334.bp.	CAA.ACA.TGC.AGgtaaggagtgt......	.6.kb..ttttcccctag	T.TGT.TAC.AGA.AGA
	...H...L...N(415)Stop.........	....polyA.........	
E7.506.bp.	CAT.CTG.AAC.TAA.GATCATACC..	ATTGTATTATTAAgctgtgaag	

Figure 17 (2 of 2)

Figure 18

Figure 19

Figure 20

3-junction:
(-12)CCCGAGGAGGATCGAGAATTAATTCCCC(12)

Aval(-12)

ATG(361)

AccI(836)

VEGF-C (2,027 bp)

TAA(1617)

293 bp
EcoRV
208 bp
HindIII

K14 poly A

5-junction:
(2020)AAAAAGCGGGCCGATCCTCTAGAG(2042)

K 14
VEGF-C
FL (EK)
~7.3 kb

Aval

EcoRI

970 bp

K 14 promoter

EcoRV

1014 bp

Reducing

Figure 21A

wt VEGF-C

Separate NR forms   IP

63
52

VEGF-C
monomer
forms

32/29

21

15

1   2   3   4   5   6   No of NR band

Non-reducing

Figure 21B

Separate NR forms   IPs

R102S   wt   VEGF-C
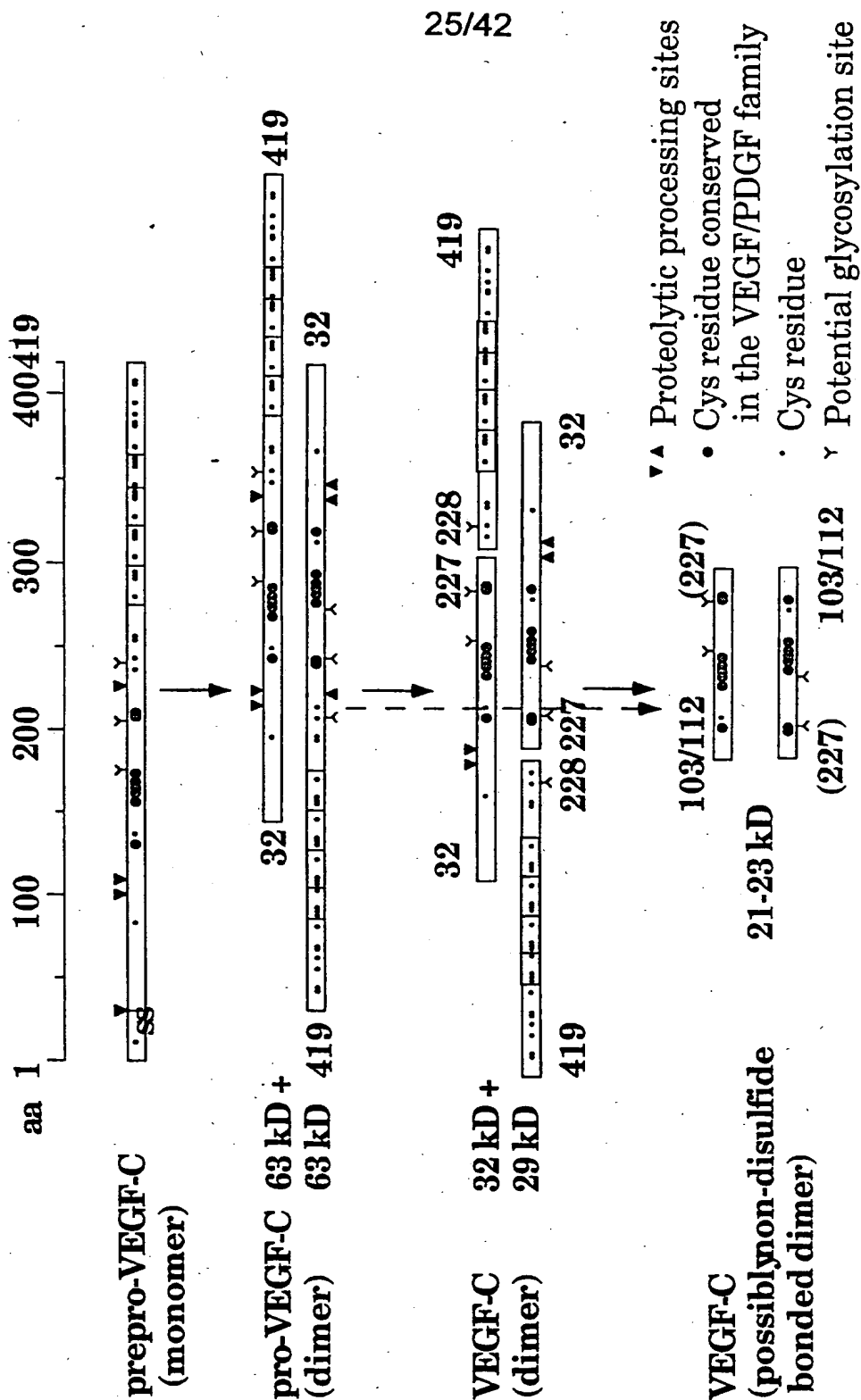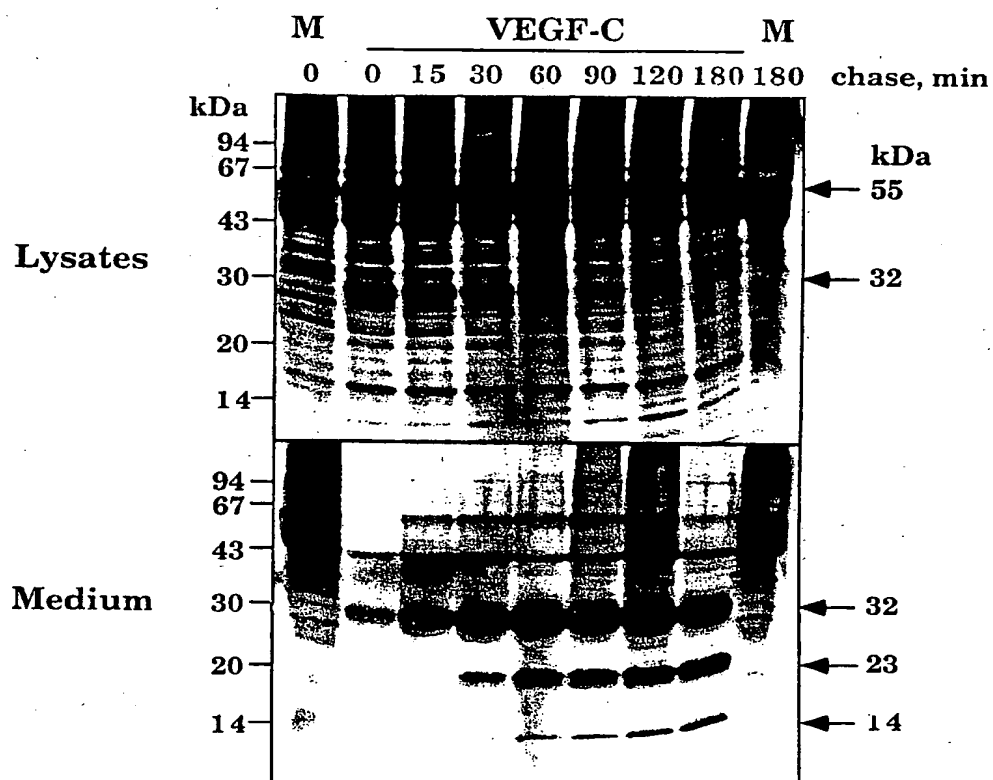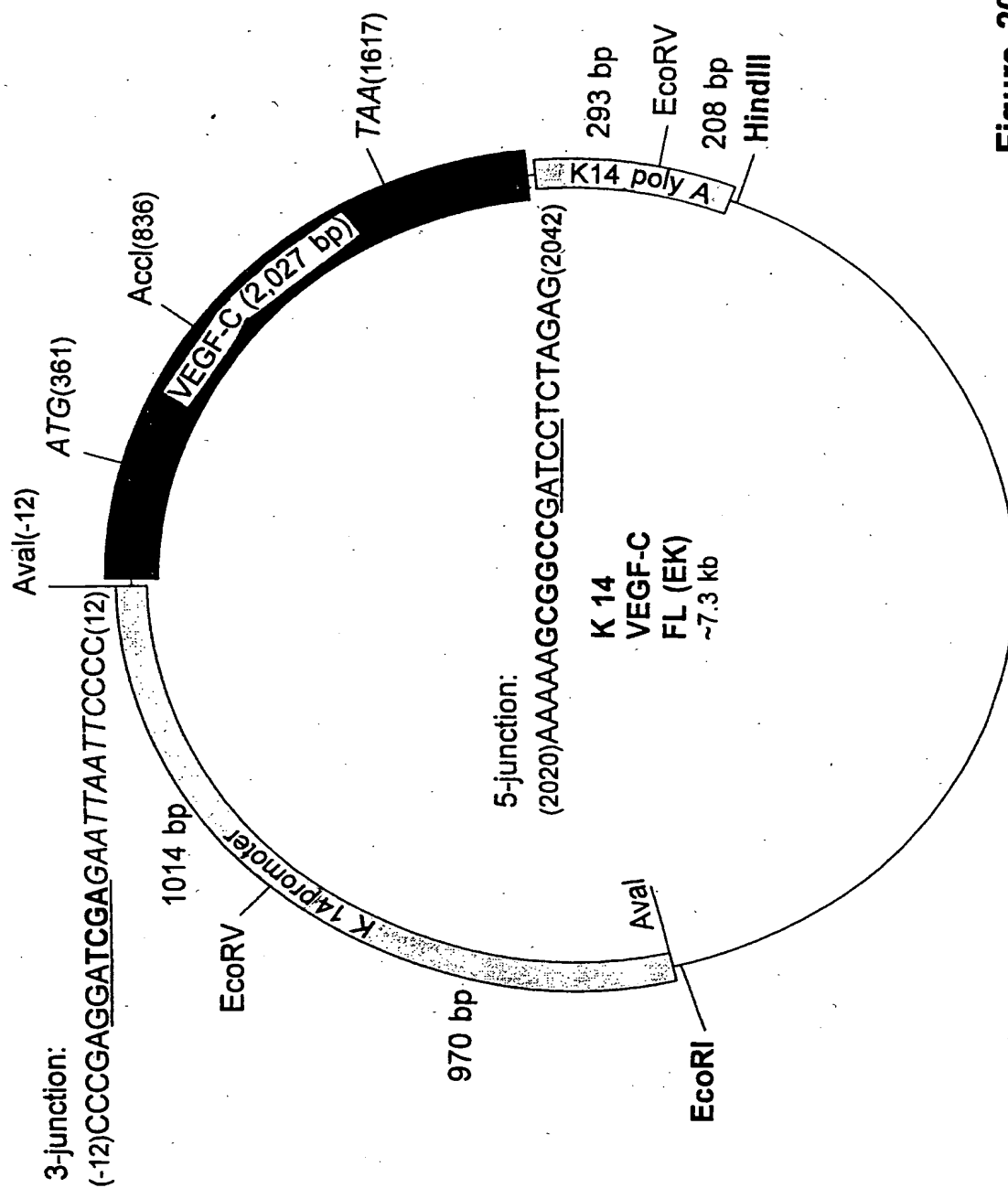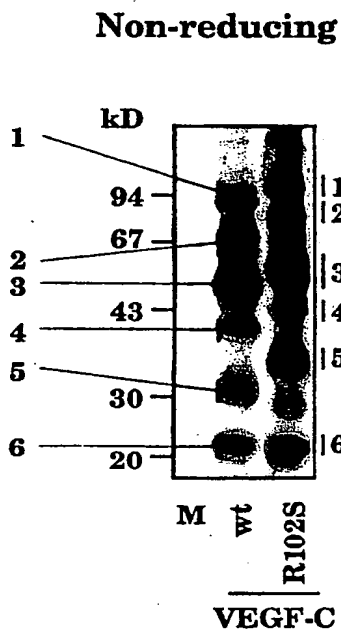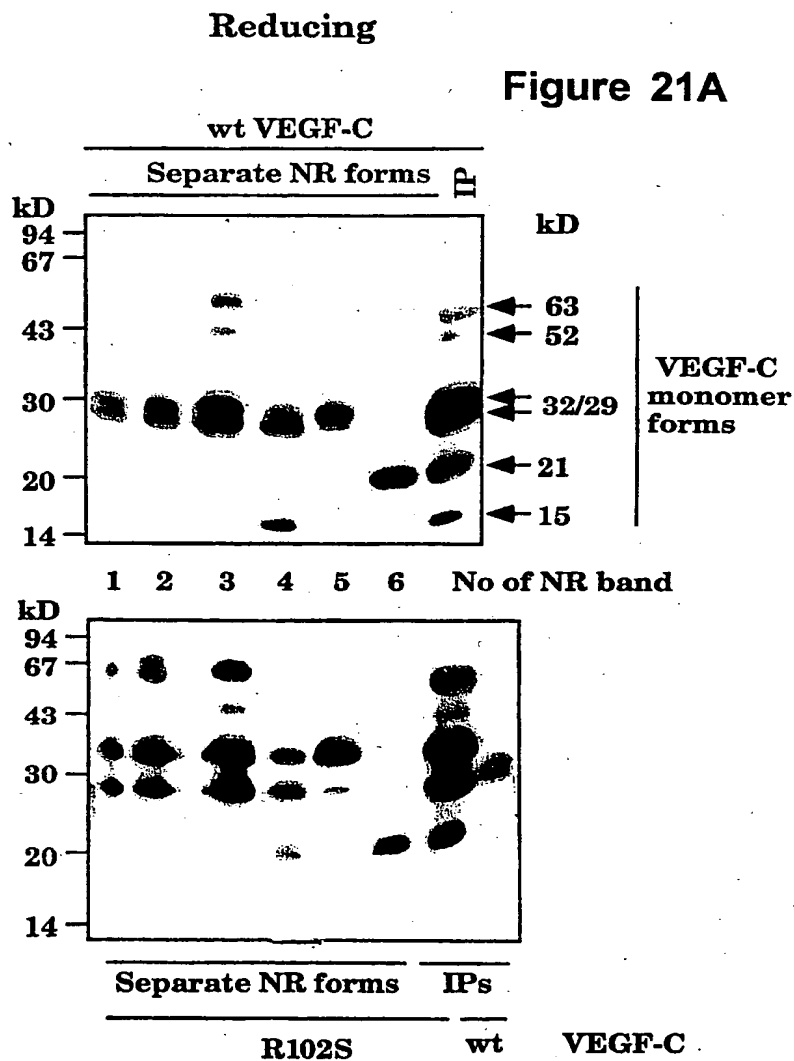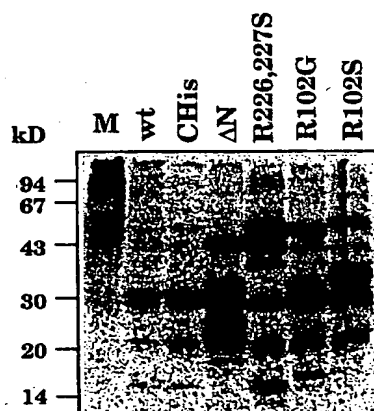
Figure 21C

Media   FIGURE 22A
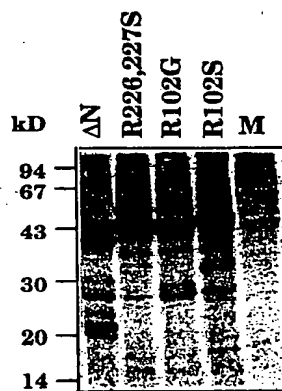
Lysates   FIGURE 22B
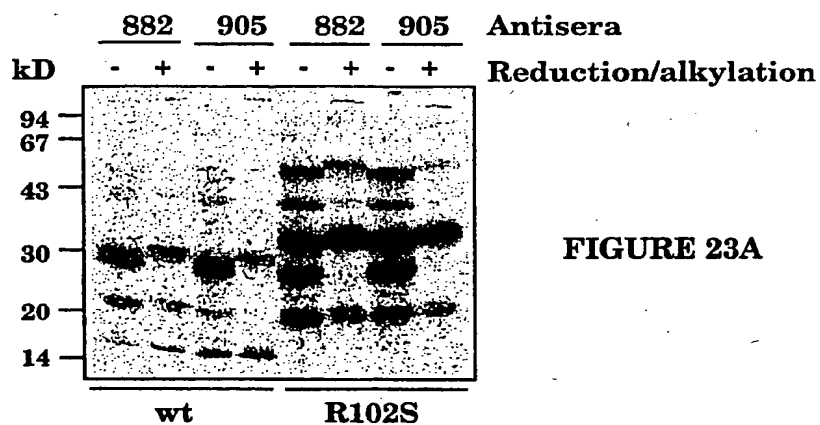
FIGURE 23A
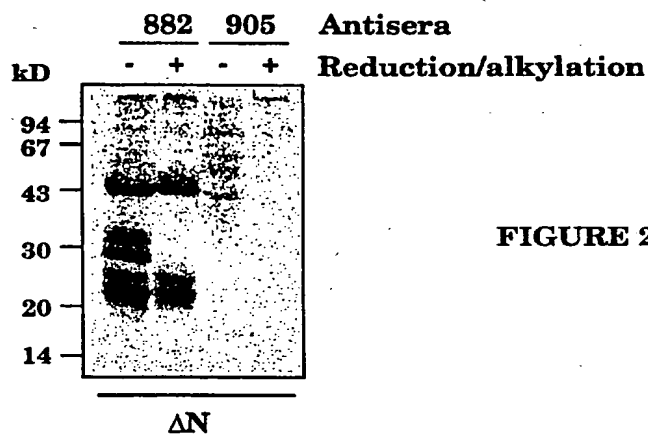
FIGURE 23B

Figure 24 A

Figure 24 B

Figure 25

induction   0 12 24   48 60   0 60 h

VEGF-C

~24 kD ⟶   anti-VEGF-C WB

M   VEGF-C   vector

Figure 26

stimulation   VO₄ - + VO₄ - + VEGF-C

195⟶   ⟵ 205

125⟶   anti-VEGFR IP
                                                anti-PTyr WB

M   Flt4   KDR
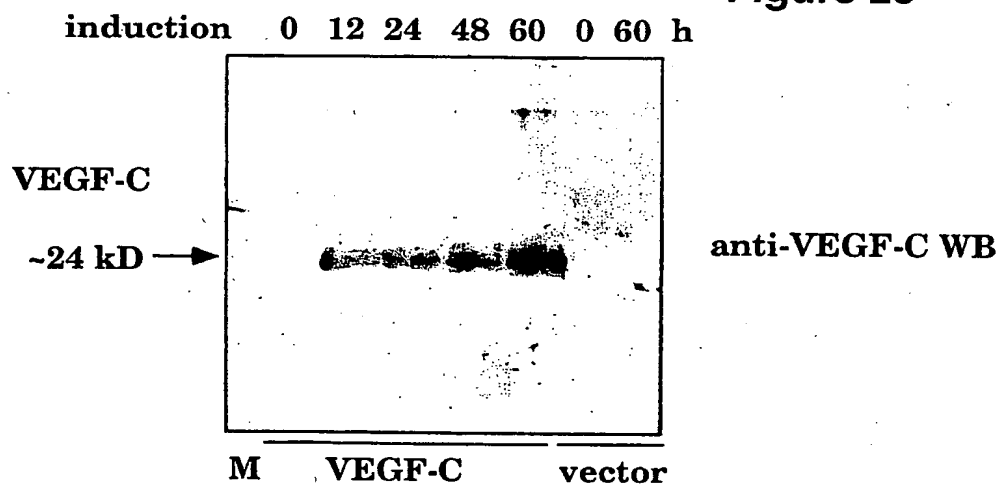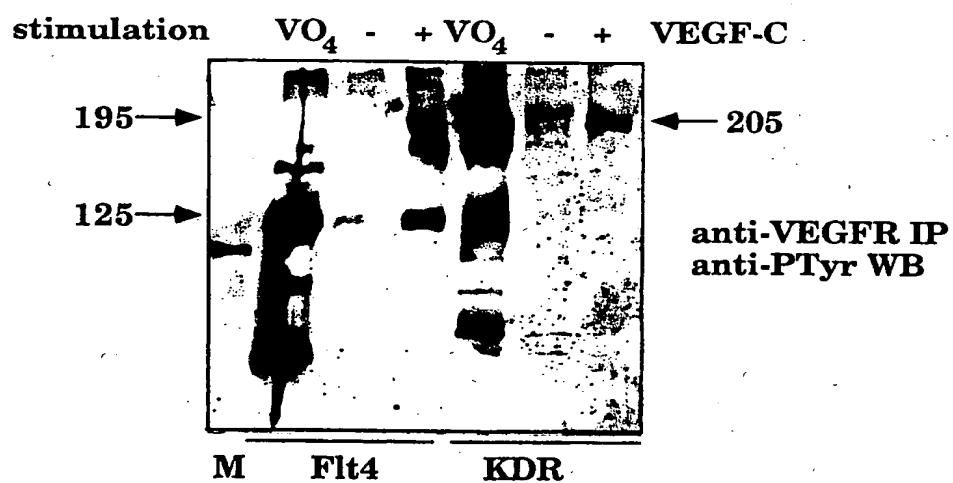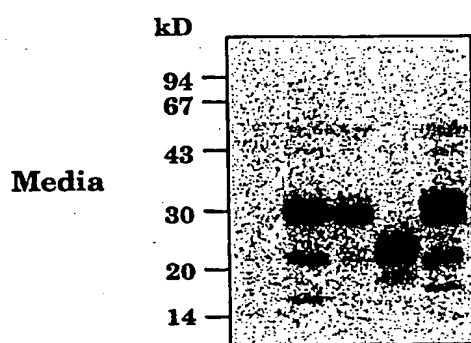
FIGURE 27A
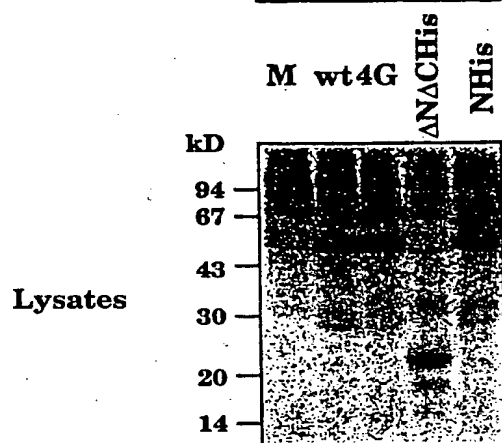
FIGURE 27B
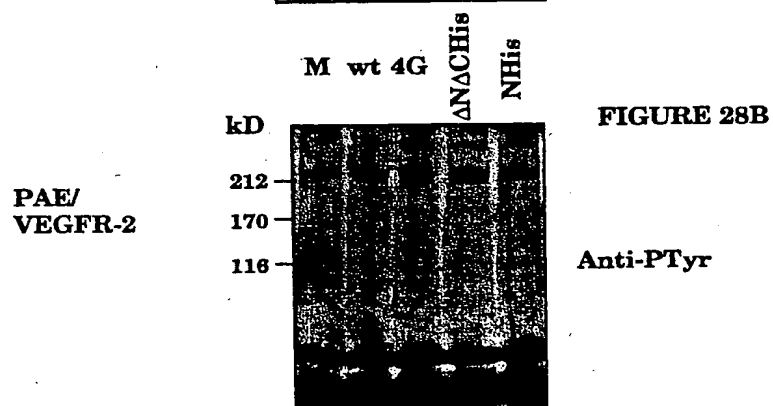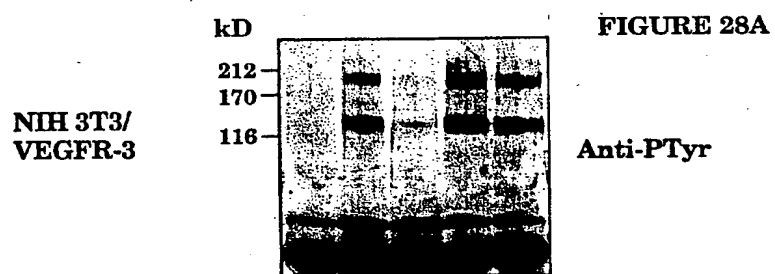
FIGURE 28A

FIGURE 28B

Figure 29A

Figure 29B

Figure 29C

Figure 29D
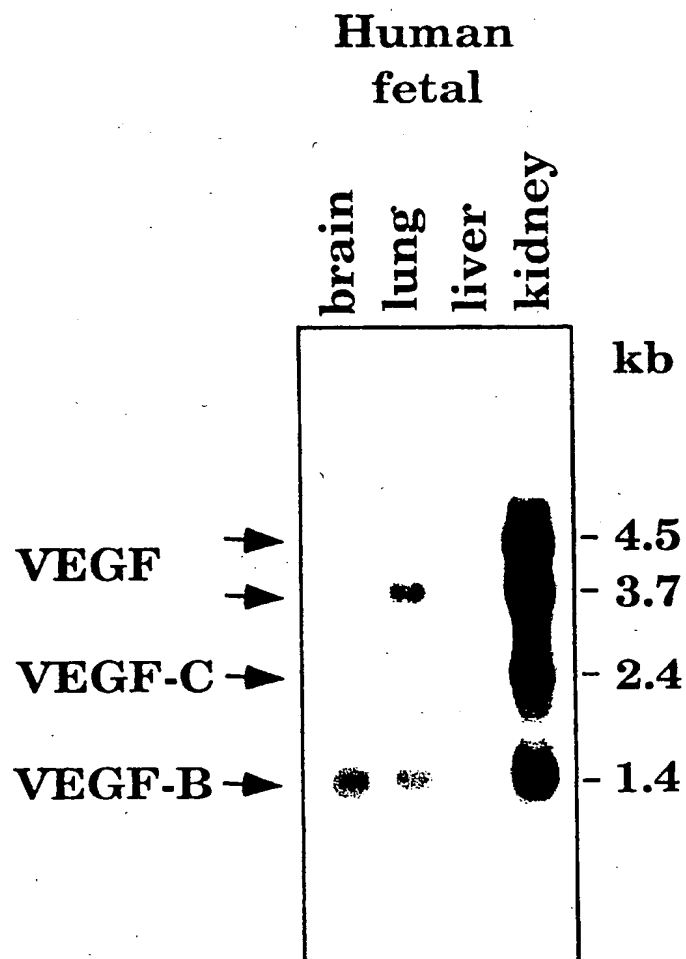
Figure 30

Signal sequence   31 ↓ 32   N-terminal propeptide

mouse   MHLLCFLSLACLLAAALIPSPREAPATVRA   FESGLGFSEAEPDGGEVRAFEGKNLEEQLRSVSSVDELMSVLYPDYWKMYKCQLRGGWQ....QPTLNTR   98

human   ...G.F.V.......L.G.......AA.   .....DL.D...A..AT.YAS.D.......T...E.........HNRE.AN..S.

VEGF homology

  C

99   ....C....RC..CC...C..............C.C.........:......   222

TGDSVKPAAAHYNTEILKSIDNEWRKTQCMPREVCIDVGKEFGAATNTFFKPPCVSVYRCGGCCNSEGLQCMYTSGYLSKTLFEITVPLSQGPKPVTISFANHTSCRCMSSKLDVYRQVHSIIR

.EETI..................................V....................S..

BR3P homology

  C.............C.C.C.........:........

223   ....

RSLPATLQCQAANKTCPTNYVWNNYM   CRC   LAQQDPIFYSNVEDDSTNGFHDV   CGPNKELDEDTCQCVC   RGGLRPSS

......................M..HI   ....E..M.S.DAG....D....I   ......E.........RA...A.

....   CGPHKELDRDSCQCVC   KNKLFPNS

....N.......SQ

....   CGANREFDENTCQCVC   KRT

....R.............

....   CPRNQPLNPGKCACEC

....

.TENTQKCFLKGKKFHHQT   CSC   YRRP

....SP..L...........   415

.CANRLKHCDPGLSFSREV   CRC   VPSYWKRPHLN

.....T..Q.A.E..GF.Y...   .QMS

Figure 31

Figure 32

Figure 33

Figure 34A   **Figure 34B**

Figure 34C   **Figure 34D**
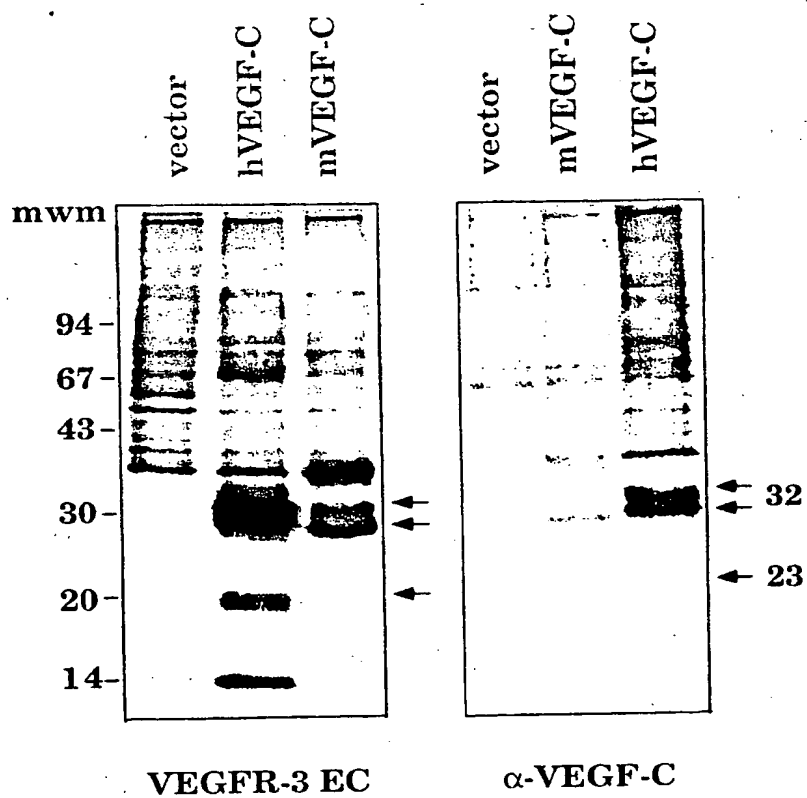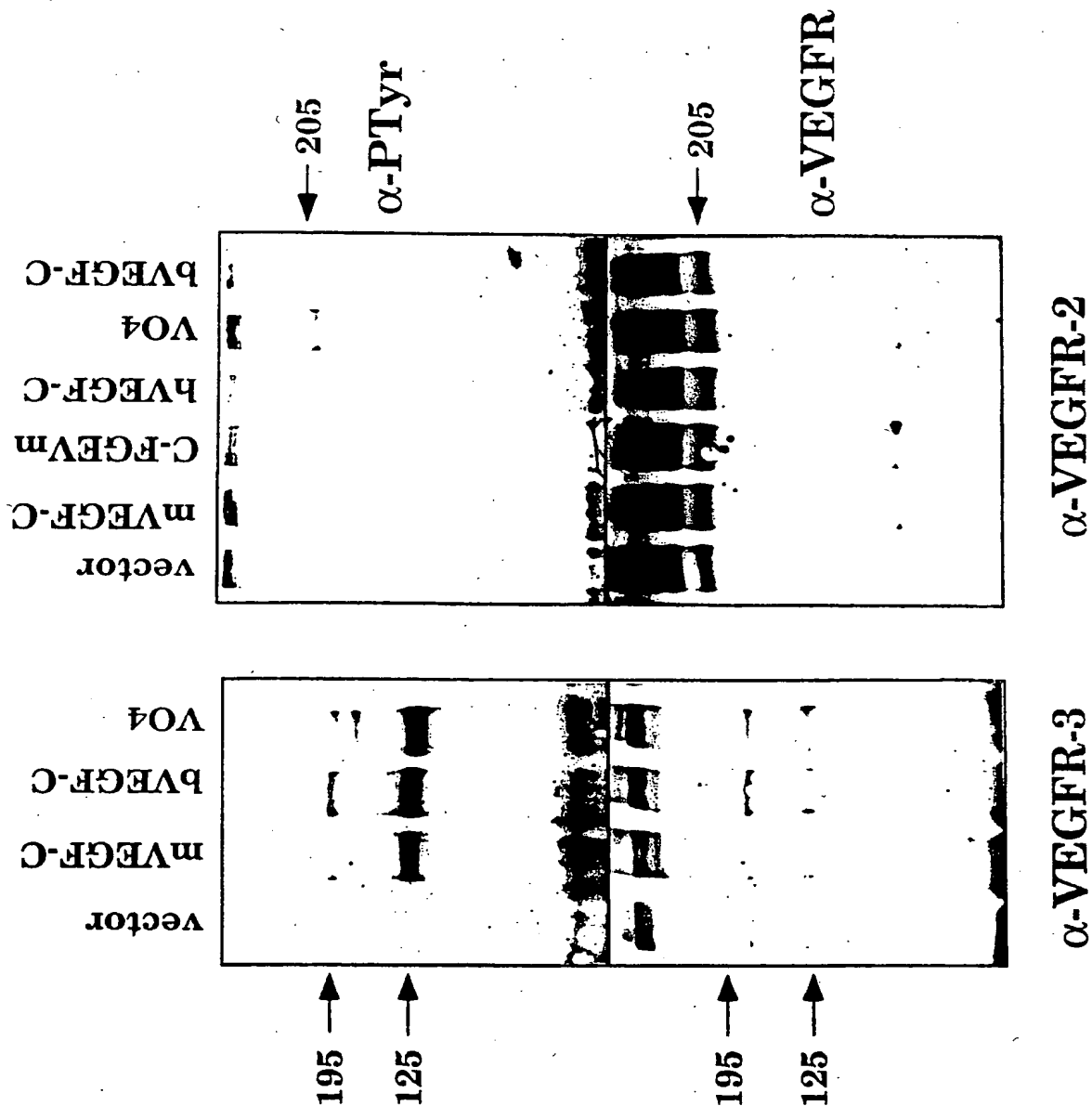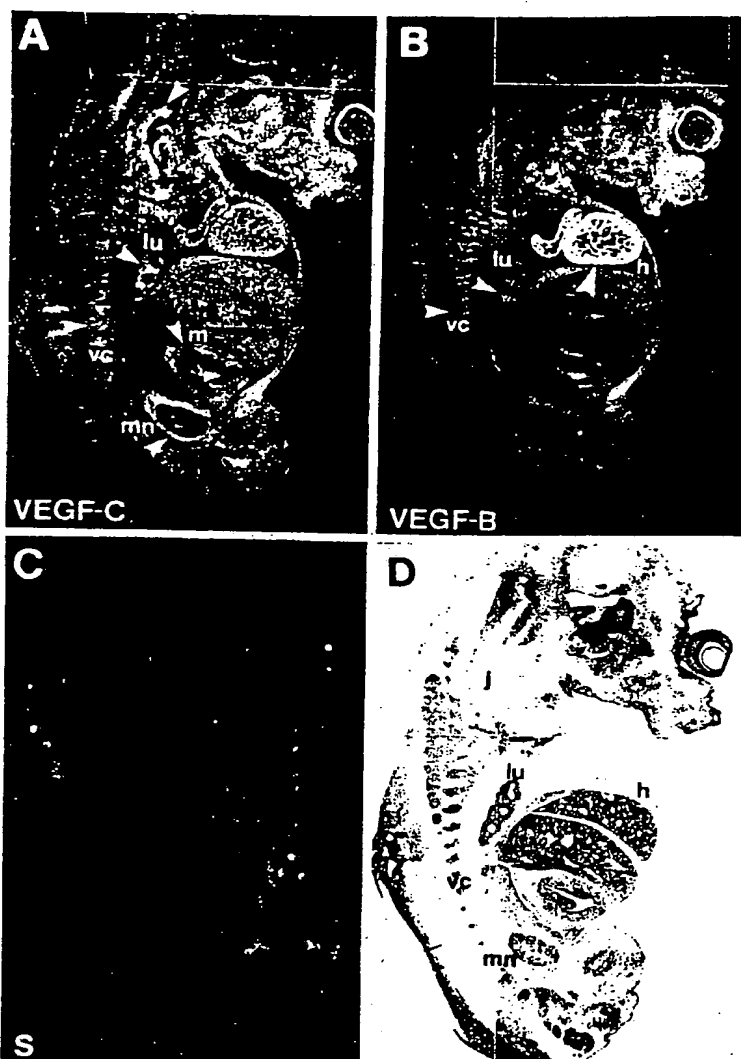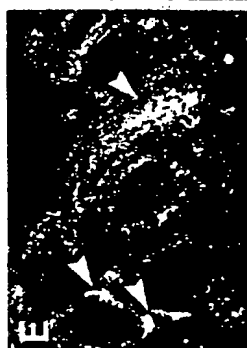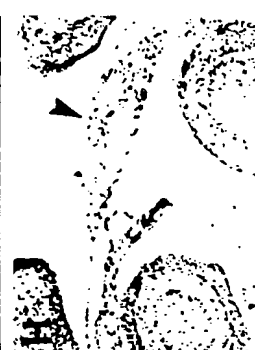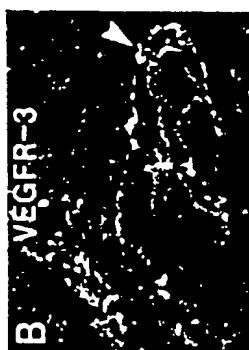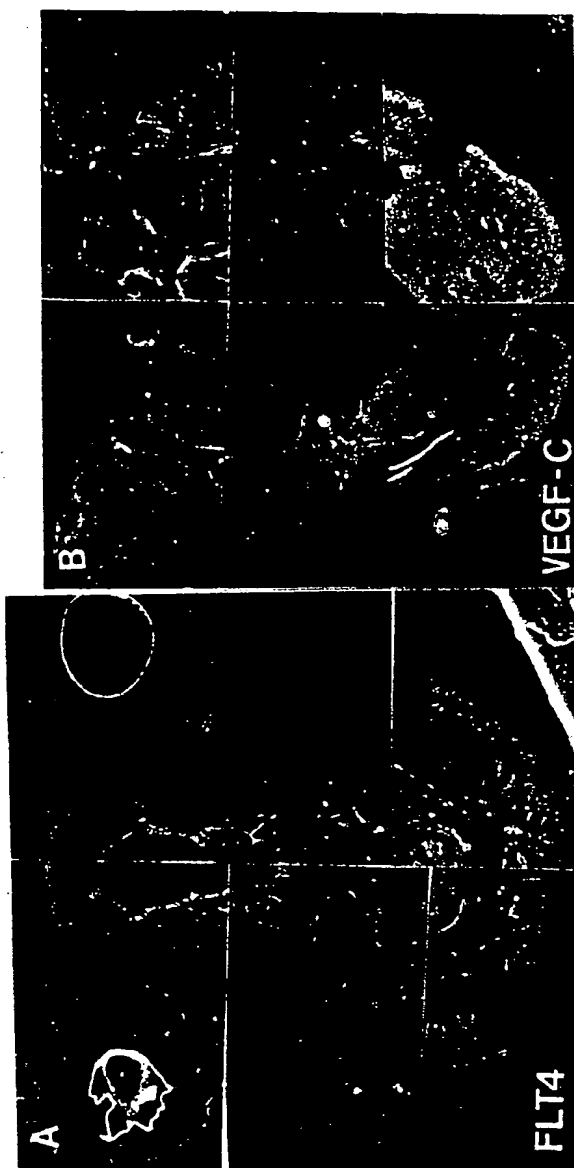
Figure 35A Figure 35B

Figure 35C Figure 35D

Figure 35E Figure 35F

Figure 35G Figure 35H

Figure
36C

Figure
36D

Figure
36A

Figure
36B